# A Statistical Framework for Model Comparison in Particle Physics

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Gareth Luke Williams

September 2019

# Acknowledgements

I would like to thank my parents for their support throughout without which this thesis would not be possible. I also want to thank my supervisors Dr Thomas Teubner and Martin Gorbahn for all their support, academic and pastoral throughout this PhD process.

# Contents

**Abstract**

With the vast quantities of data from experimental work an important task will be to distinguish which physical models are best fitted to these data sets especially in cases where the data shows deviations from the Standard Model. The Bayes factor is a tool used to determine which of two competing hypotheses is favoured given the data and in the case of model comparison provides the ability to determine which of the two physical model is favoured subject to the data. The marginalized likelihood is obtained by integrating the product of the likelihood $P(D|\vec{\theta}_i)$ and the prior distribution of the parameters $P(\vec{\theta}_1|M_i)$ over the parameter vectors $\theta_i$ for some model $M_i$ given data D. Calculating the marginalized likelihood of two models and taking the ratio one obtains the Bayes factor. The degree to which one model is favoured over the other is obtained from a standard table which offers an interpretation of the result depending on which band of values the resultant Bayes factor lies.

The computational framework to allow one to undertake statistical analysis of this kind is presented with a discussion on how it was built along with the underlying statistics to allow a user to input any model desired that they may wish to perform such analyses on.

The minimally supersymmetric Standard Model and two-Higgs doublet model represent two of the simplest extensions to the Standard Model which both offer explanations to some of its deficiencies. Following the discussion on the fundamentals of these two models an explanation of how they handle the anomalous magnetic moment of the muon and lepton flavour violating decays is given and the Bayes factor calculated between these two models. The result of this indicates the two-Higgs doublet model is preferred given the data on these observables and the reasoning behind why such a result is obtained is presented in the final chapter of this thesis.

# Chapter 1

# Introduction

The Large Hadron Collider (LHC) has yielded great success with the myriad of results produced. The discovery of the Higgs boson in 2012 at the mass region of approximately 126 GeV meant that the particles outlined in the Standard Model had finally all been experimentally observed. This added credence to the Standard Model, which is one of the most well tested physical theories predicting a broad range of phenomena and successfully explaining almost all experimental outcomes. Grand Unification Theories (GUT) in the Standard Model are successful in unifying three of the four fundamental forces at high energies, the electromagnetic, the weak and strong nuclear forces. This failure of the Standard Model to unify gravity with the other forces is perhaps its most well known failure and therefore serves as a major indicator that the Standard Model is perhaps an incomplete picture of particle physics.

Another indicator of physics beyond the Standard Model is the existence of neutrino oscillations first theorised to exist in 1957 by the Italian physicist Bruno Pontecorvo in his paper [1] and later expanded upon in [2]. Neutrino oscillations were unknowingly observed experimentally in the 1960s [3] which was known as the solar neutrino problem and in 2001 the cause of this problem was indeed identified as neutrino oscillations [4]. However the first demonstrable evidence was provided in 1998 [5] by Super-Kamiokande Collaboration and provided an experimental result that also indicated a deficiency in the standard model since neutrino oscillations require neutrinos to have non-zero mass, a requirement that is in total disagreement with the standard model prediction of massless neutrinos.

The fact that there is an observed imbalance of matter to antimatter in the universe is so far also unexplained by the Standard Model. The conditions for this so called matter-asymmetry were given by Sakarov in 1967 [6]:

- Violation of baryon number,
- Violation of C- and CP-symmetry,
- The interaction of baryons producing matter and anti-matter at different rates must go out of thermal equilibrium.

The first of these Sakarov conditions could possibly be answered by sphalerons, a static solution to the electroweak field equations in the Standard Model. It is predicted that sphalerons would be observed around the 10 TeV level and although current runs at the LHC are at the 13 TeV level, this energy cannot be concentrated specifically to generate sphalerons and as a result they are yet to be observed. Although CP-symmetry violation has been observed in neutral Kaon decay in 1964 [7], there is observed no mechanism in the standard model to allow for baryon number violation. These conditions give indications to the possible need of beyond the standard model physics.

Dark matter and dark energy are both popular topics in science fiction however they indeed present a very real conundrum in reality. The observed rotation of galaxies in addition to their calculated mass [8], indicates the existence of large quantities of matter which has so far not been directly detected. The discussed violation of Kepler's second law, which proposes the velocity of objects in the galaxy decreases the further from the centre the object is, along with the disagreement between theory and experiment of the predicted velocity distribution of elliptical galaxies and the mass of galaxy clusters being observed far greater than predicted give predictions for dark matter to be far more prevalent than visible matter on a scale of $5:1$ [9]. Dark energy is a hypothesized form of energy that is responsible for the accelerating expansion of the universe [10]. With dark energy thought to contribute 68% of energy in the universe and dark matter outweighing visible matter on the scale discussed, the standard model gains yet another major indicator to its deficiency.

With the shortcomings discussed above and with the standard model containing at least 28 arbitrary parameters [11] it is important to consider possible models beyond the Standard Model given the immense quantity of experiments

currently underway to investigate these problems. An important task with the continuation of experimental work is to examine physical models that may be able to compensate for the shortcomings of the Standard Model. To do this one must devise a statistical method to distinguish between these models, one such method is to calculate the Bayes factor between two models which produces a standardised quantity whose value can be matched to the established list of possible values indicating which model is preferred and how strongly.

To produce such an output a computational framework was devised which was written in Python, the structure of which is explained in greater depth in chapter 4 and later. It was a crucial component to the analysis to ensure that the framework produced correct results so to ensure rigour in this project test cases were selected whose output was known previously. Once the output was compared with results previously established the framework could be extended to include observables not previously analysed in this manner in different physics models to allow the Bayes factor to be calculated and achieve the aim of the project to produce the desired statistically valid comparison between models.

The observables examined here were the anomalous magnetic moment of the muon, $a_\mu$ and lepton flavour violating observables, including the branching ratio $Br(\mu \to e\gamma)$. The vast amount of current experimental and theoretical work being undertaken and their link through the similarity of the Feynman diagrams made these observables an interesting topic of research with regard to statistical comparison. Chapter 3 delves into the reasoning and motivation behind the examination of the observables in more detail.

It was decided to consider the Minimal Supersymmetric Standard Model (MSSM) and the two-Higgs doublet model (2HDM) since these are two of the most well known and studied models that also represent two of the simplest extensions to the standard model. Later chapters discuss the MSSM and 2HDM in more detail and discuss the theoretical underpinnings that make them interesting to study regarding searches for new physics and how they offer possible explanations for the dependency in the measurement of the anomalous magnetic moment of the muon while also offering the ability to allow for lepton flavour violating decays.

# Chapter 2

# Statistics Overview

This chapter discusses the two main approaches to statistical inference, frequentist and Bayesian. Then follows a discussion on why Bayesian statistics offers the desired approach for statistical inference and then the tools offered by this approach for model comparison are set out, specifically the Bayes factor which provides a method for comparing two hypothesis (in our case two physical models) and by the processes laid out produces a numerical quantity that one can use to classify how well preferred one model is over the other. Finally in the chapter is the discussion on how the likelihood function is constructed and generalised to n-dimensions by introducing the covariance matrix and how experimental data is input into this.

## 2.1 Probability Axioms

Statistical data analysis is mainly focussed on fitting (and comparing) models to a measured data set. The term inference is a term commonly used in statistics and is used to describe one deciphering meaning from data.

Two schools of inference are frequentist and Bayesian statistics, each having different axioms of probability. For the frequentist approach to statistics the Kolmogorov axioms are used to formalise probability and are [12],

- The probability of an event is a real, non-negative number,
  $P(E) \in \mathbb{R}$, $P \geqslant 0$, $\forall E \in S$,
  where $S$ is the event space.

- An elementary event is an event with only one outcome and the proba-

bility of at least one such event occurring in the entire sample space is one,
$P(\Omega) = 1$.

- Any countable sequence of disjoint sets $E_1, E_2, ...$ satisfies
$P(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$.

For the Bayesian approach to statistics, Cox's theorem is largely favoured for this task. The three components to the theorem are [13],

- The probability of an event is a real number with a dependency on known information of the event,
$P(E) \in \mathbb{R}|I$,
where $P(E)$ denotes the probability of the event and $I$ represents the known information.

- If there are multiple ways to derive the probability of an event then all of these methods must produce the same outcome.

- Probabilities can vary after prudent assessment of probabilities in the model.

## 2.2 Why Bayesian?

The interpretation of experimental results can vary on whether one takes the approach of a Bayesian or frequentest statistician; Bayesian statistics may be more appealing depending on the context. This school of statistics has one think about what one is asking of the data, how does the data affect one's prior knowledge of the experiment?

A major difference between the two schools of statistical thought is in their treatment of the likelihood function. In the classic frequentest school it is used to approximate confidence intervals, but the likelihood function has a more fundamental role in Bayesian statistics. Bayesian statistics differs in its treatment of the likelihood function by integrating it and using it to calculate a probability density function [15] and by its having a prior probability function and a posterior probability and using them with Bayes' theorem as outlined previously and integrating to obtain the marginalized likelihood.

The main advantage of the Bayesian approach over the frequentist approach is that of hypothesis testing. In classical frequentist statistics one is concerned with testing some hypothesis $H_1$ against a null hypothesis $H_0$ which does not allow one to easily compare models as desired here, whereas Bayesian allows the testing of two hypothesis relative to each other and is the approach used here to calculate the Bayes factor to measure the favourability of one model over another.

## 2.3 Bayes Factor

Bayes' theorem forms the basis of Bayesian statistics and has the form,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \tag{2.1}$$

where $P(B|A)$ is a conditional probability, the probability of $A$ given $B$. The purpose of this project is to compare physical models in a statistically meaningful manner. To do this, we consider a hypothesis that some physical model, $M$ is true and some experimental or theoretical data $D$ to test the model and write Bayes' theorem as,

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}. \tag{2.2}$$

where $P(M|D)$ is the posterior probability and represents the conditional probability of the model after the data is considered and $P(D|M)$ is the probability that the data is produced according to the model, $M$. Here $P(M)$ represents the probability of the model before any of the data is taken into account and is usually referred to as the *prior* probability. The likelihood $P(D|M)$ denotes the probability that data $D$ is produced according to model $M$. It is the important part of equation 2.2 for our analysis because the Bayes factor is defined as the ratio of these likelihoods of two models,

$$B_{12} = \frac{P(D|M_1)}{P(D|M_2)}. \tag{2.3}$$

The likelihoods in equation 2.3 are known as *marginalized* likelihoods. Given a model $M_i$ with observables parameterized by the vectors $\vec{\theta}_i$, the marginalized likelihood is the likelihood after the parameter vectors have been integrated out,

$$B_{12} = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\int_{\vec{\theta}_1} P(D|\vec{\theta}_1)P(\vec{\theta}_1|M_1)d\vec{\theta}_1}{\int_{\vec{\theta}_2} P(D|\vec{\theta}_2)P(\vec{\theta}_2|M_2)d\vec{\theta}_2}. \tag{2.4}$$

where $P(D|\vec{\theta}_i)$ is the joint probability of the data given the parameters of the model and $P(\vec{\theta}_i|M_i)$ is the distribution of the parameters $\vec{\theta}_i$ in the model $M_i$ and it represents the prior probability of the parameters. These quantities on the right hand side of equation 2.4 are calculable in our code allowing us to then compute the Bayes factor.

The Bayes factor can be used to test evidence for some model, $M_2$, against another model $M_1$. The integration can by done in Python using the SciPy package, however after the sampling is complete we are left with a discrete list of data points meaning the Bayes factor can be calculated by summation rather than integration,

$$B_{12} = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\sum_{\vec{\theta}_1} P(D|\vec{\theta}_1)P(\vec{\theta}_1|M_1)}{\sum_{\vec{\theta}_2} P(D|\vec{\theta}_2)P(\vec{\theta}_2|M_2)} \tag{2.5}$$

This calculation was performed by taking the list of the values for the likelihoods $P(D|\vec{\theta}_i)$ and and for each element of the list, multiplying it by the prior $P(\vec{\theta}_i|M_i)$,

```
while i<len(likelihood):
        bfact=likelihood[i]*prior[i]
        bayeslist.append(bfact)
        i=i+1
bayesfactor=np.sum(bayeslist)
```

## 2.4 Bayes Factor Interpretation

The interpretation of the values was given by Jeffreys (1961) can be used to determine how strongly supported one of the models is by the given data [14],

This table gives a simple method of interpreting the output to determine which model is favoured and to what degree.

| $log_{10}B_{12}$ | $B_{12}$ | Evidence against $M_2$ |
|---|---|---|
| $< 0$ | $< 1$ | Evidence favours $M_2$ or evidence is against $M_1$ |
| $0 - \frac{1}{2}$ | $1 - 3.2$ | Barely any |
| $\frac{1}{2} - 1$ | $3.2 - 10$ | Substantial |
| $1 - 2$ | $10 - 100$ | Strong |
| $> 2$ | $> 100$ | Decisive |

**Table 2.1:** *Table showing the possible value for the Bayes factor $B_{12}$ between model $M_1$ and $M_2$ and the interpretation of this value in the third column*

## 2.5 The Covariance Matrix

Experimental results of observable measurements are usually reported in the form,

$$\mu \pm \sigma. \tag{2.6}$$

where $\mu$ is the mean value and $\sigma$ is the standard deviation. So for example, $a_\mu = (26.8 \pm 7.6) \times 10^{-10}$ then $\mu = 26.8 \times 10^{-10}$ and $\sigma = 7.6 \times 10^{-10}$. A normally distributed likelihood for a one dimensional quantity has the form,

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.7}$$

However since the aim is to build as general a framework as possible it is sensible to include the ability for the user to calculate likelihoods of models that contain more than one observable. To do this the form of the likelihood must be extended to $n-$dimensions which is achieved by not only considering the mean and measured values as vectors but also instead of using the standard deviation as the well known component of the normal distribution one uses the covariance matrix of the observables in its place.

Often observables are discussed in terms of being correlated or uncorrelated. To see the impact this has on calculations first let there be $n$ random variables, $X_1, X_2, ..., X_n$. The correlation matrix is defined as an $n \times n$ matrix where the $i, j^{th}$ entry is given by $corr(X_i, X_j)$, where

$$corr(X_i, X_j) = \rho_{X_i X_j} = \frac{\sigma_{X_i X_j}}{\sigma_{X_i}\sigma_{X_j}} = \frac{cov(X_i, X_j)}{\sqrt{Var(X_i)}\sqrt{Var(X_j)}}. \tag{2.8}$$

The correlation coefficient defined above takes values in $[-1, +1]$ where a positive value indicates that as one variable increases, so does the other and a

negative value means that one variable increases as the other decreases [16]. The correlation matrix denoted by $\bar{\rho}$ is composed of the correlation coefficients of the observables and is given by,

$$\bar{\rho} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & & & \\ \rho_{31} & & 1 & & \\ \vdots & & & \ddots & \\ \rho_{n1} & & & & 1 \end{bmatrix}$$

In the case where we have independent variables the correlation coefficient is zero, so we end up with a diagonal correlation matrix,

$$\bar{\rho} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & & & \\ 0 & & 1 & & \\ \vdots & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} = \mathbb{I}_{\mathrm{n\times n}}$$

.

When discussing multiple observables one often talks in terms of whether or not the observables are correlated however when one wishes to perform calculations involving the observables it is often more convenient to use the covariance matrix rather than the correlation matrix. This is achieved by using equation 2.8 to convert the correlation matrix into the covariance matrix denoted $\Sigma$,

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & & & \\ \sigma_{31} & & \sigma_3^2 & & \\ \vdots & & & \ddots & \\ \sigma_{n1} & & & & \sigma_n^2 \end{bmatrix}$$

In the case of uncorrelated observables, the covariance matrix is simply composed of the standard deviations of the observables along the diagonal and

all other entries as zero.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & & \\ 0 & & \sigma_3^2 & & \\ \vdots & & & \ddots & \\ 0 & & & & \sigma_n^2 \end{bmatrix}$$

This then allows one to extend equation 2.7 to an n-dimensional Gaussian which has the form,

$$P(\vec{x}) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} e^{\left(-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T\right)}, \tag{2.9}$$

where $\vec{x}$ and $\vec{\mu}$ denote vectors of the calculated observable values and mean values respectively. $\Sigma$ denotes the covariance matrix, $|\Sigma|$ is its determinant. It is important to note that in our notation that $\vec{x}$ and $\vec{\mu}$ are row vectors so the transpose is to the right of the $\Sigma^{-1}$ term for the matrix multiplication to work. Equation 2.9 then gives the desired expression required for the calculation in the project to be then used in the calculation of the Bayes factor.

One can see that for model comparison Bayesian statistics has a very useful tool in that of the Bayes factor which is calculated as a the ratio of marginalized likelihoods for two models. The marginalized likelihood is shown to be calculated by summing over the product of the likelihood $P(D|\vec{\theta_i})$ and the distribution of $\vec{\theta_i}$ in the model $M_i$, $P(\vec{\theta_i}|M_i)$ also called the prior probability. The likelihood can be calculated in n-dimensions by using equation 2.9 with the construction of vectors of the mean values and sampled values along with the covariance matrix of the observables. Performing this calculation in two different models and taking the ratio of the results gives a value which one can easily use table 2.1 to interpret the result and conclude which model is favoured and how strongly it is favoured.

# Chapter 3

# Observables Overview

To compare two models it is important to determine which observables to consider in each model. The Bayes factor is defined as the ratio of likelihoods $P(D|M_i)$ so it is important when comparing models in this manner to consider observables with sufficient experimental data. The current and future experimental work outlined in this chapter regarding the anomalous magnetic moment of the muon and lepton flavour violating decays allows one to consider these observables in models one wishes to compare in this manner. The use of beyond the Standard Model physics as an attempt to explain these discrepancies also made them an attractive candidate for statistical analysis involving model comparison. An introduction to the anomalous magnetic moment of the muon $a_\mu$ is presented with an overview of how it is calculated in the Standard Model and how experiments measure this value and show a discrepancy. The similarity between the Feynman diagrams of $a_\mu$ and the lepton flavour violating (LFV) decay $\mu \to e\gamma$ is discussed which, along with the interest of LFV regarding physics beyond the Standard Model, makes them interesting to consider together in our analysis.

## 3.1 $a_\mu$ Overview in the Standard Model

The Standard Model of particle physics has had remarkable success in providing experimental predictions however it does have shortcomings, one of which is currently the anomalous magnetic moment of the muon. The magnetic moment for an elementary charged particle is,

$$\vec{\mu} = g_\mu \frac{e}{2m} \vec{s}. \tag{3.1}$$

Where $g_\mu$ is the gyromagnetic factor and has a predicted value $g_\mu = 2$ from the Dirac equation, $e$ is the elementary electric charge and has a value of 0.3028 in natural units $\hbar = c = 1$ and $\vec{s}$ is the spin vector. There is a deviation from the predicted $g_\mu = 2$ which arise from quantum loop corrections parametrized as,

$$a_\mu = \frac{g_\mu - 2}{2} \tag{3.2}$$

hence the commonly used name '$g-2$'.

In the Standard Model the anomalous magnetic moment of the muon can be separated into the following constituents [17],

$$a_\mu^{SM} = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{Hadronic} \tag{3.3}$$

where $a_\mu^{QED}, a_\mu^{EW}$ and $a_\mu^{Hadronic}$ denote the quantum electrodynamic, electroweak, and hadronic contributions to the anomalous magnetic moment of the muon, see Fig. 3.1 for examples of diagrams of these components.

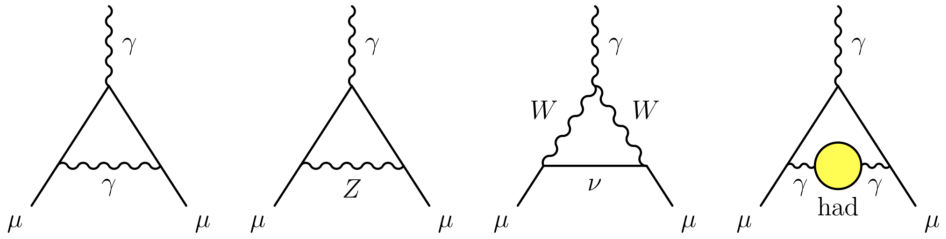



**Figure 3.1:** *Feynman diagrams from Ref. [17] representing the Standard Model constituents of $a_\mu$. The diagrams are the first order QED, the lowest order weak and lowest order hadronic components respectively.*

The calculation of $a_\mu^{SM}$ has been the task of numerous analyses with the difference between this value and the experimentally measured value calculated as [17],

$$a_\mu^{EXP} - a_\mu^{SM} = (26.8 \pm 7.6) \times 10^{-10} \tag{3.4}$$

There is work being undertaken to improve the accuracy of the Standard Model calculation for $a_\mu$ with the aim of eliminating the difference in equation 3.4. However this difference could also be an indicator of physics beyond the Standard Model which is why this observable was one of the main focuses of this project.

Experiment E989 offers a measurement the anomalous magnetic moment of the muon by using the cyclotron frequency of the storage rings where the muons are injected [21],

$$\omega_c = \frac{e}{m\gamma}B, \tag{3.5}$$

where $\omega_c, e, m, \gamma, B$ are the cyclotron frequency, elementary electric charge, muon mass, and magic momentum respectively. The stored muons precess in the magnetic field with frequency,

$$\omega_s = \frac{e}{m\gamma}B(1 + \gamma a_\mu), \tag{3.6}$$

and the difference between these frequencies is measured giving,

$$\omega_a = \omega_s - \omega_c = \frac{e}{m}a_\mu B. \tag{3.7}$$

$$\vec{\omega_a} = \vec{\omega_S} - \vec{\omega_C} = -\frac{Q}{m}\left[a_\mu B - \left(a_\mu - \frac{1}{\gamma^2 - 1}\right)\frac{\vec{\beta} \times \vec{E}}{c}\right] \tag{3.8}$$

where $Q$ is the electric charge and the magic momentum, $\gamma = \frac{m}{\sqrt{a_\mu}} \approx 3.09$ GeV. This then allows for the measurement of the quantity $a_\mu$.

## 3.2 Lepton Flavour Violation

In the Standard Model lepton number is a conserved quantity, however neutrino oscillations gave the first indication that this is not necessarily the case. Electron and muon neutrino production is dominated by the flavour conserving muon decay, $\mu^+ \to e^+ + \bar{\nu_\mu} + \mu_e$ and the pion decay, $\pi^+ \to \mu^+\nu_\mu$ [22]. However observations from cosmic rays colliding with nuclei in the atmosphere produce hadronic showers whose decays produce atmospheric neutrinos which have shown be flavour oscillating $\nu_\mu \leftrightarrow \nu_\tau$. Neutrino oscillations offer an experimentally confirmed result of a lepton flavour violating process forbidden in the Standard Model.

The Minimal Supersymmetric Standard Model (MSSM) is an extension to the Standard Model discussed in more detail in chapter 6 which posits the existence of so called super-partners of the bosons and fermions and offers possible solutions to shortcomings of the Standard Model. Examples of short-


**Figure 3.2:** *Mass-insertion Ferynman diagrams contributing to $a_\mu$ (left) and $a_{\mu e\gamma}$ (right) involving higgsinos $\tilde{H}_{u,d}$ and winos $\tilde{W}$ from [23]*

comings possibly resolved by the MSSM are the anomalous magnetic moment of the muons $a_\mu$ and Lepton Flavour Violating (LFV) decays. In the MSSM the Feynman diagrams for $Br(\mu \to e\gamma)$ are the same as for $a_\mu$ except for the inclusion of flavour transition [23] (see figure 3.2). The similarity of the Feynman diagrams makes these two observables an interesting study in the MSSM. In the Standard Model the branching ratio $Br(\mu \to e\gamma)$ is calculated to have a immeasurably low value however given recent measurements from experiments such as Mu to E Gamma (MEG), this is known not to be the case, indicating any measurement of the $\mu \to e\gamma$ process would be suggestive of physics beyond the Standard Model. The current limit on the $\mu \to e\gamma$ branching ratio was published by MEG as $Br(\mu \to e\gamma) < 4.2 \times 10^{-13}$ [24]. Currently work is being done to upgrade the MEG experiment known as MEGII with the aim of increasing the sensitivity for the branching ratio of $\mu \to e\gamma$ to the magnitude of $6 \times 10^{-14}$ [24]. In addition to MEG there are more experiments planned to study lepton flavour violating decays such as Mu2e and Mu3e with data collection expected to start in 2020.

Although the prospect of lepton number not being always conserved is an interesting topic in of itself, for the purpose of the project here it was important to select observables to study where there would be adequate experimental data to work with when calculating the relevant conditional probabilities since these depend on experimental data. The investments into experiments of this type demonstrate an interest in the scientific community which makes LFV observables an appealing subject to statistical analysis of the type laid out in this thesis.

The study of the anomalous magnetic moment of the muon is a subject

known to be undertaken by many on the theoretical and experimental side. This is demonstrated as discussed above by the quantity of interest with work undertaken at major experiments such as E989 at Fermilab which makes it an important observable to be considered for the kind of statistical analysis undertaken in this project. Lepton flavour violating decays are also discussed above as observables with considerable current theoretical and experimental interest and their link with $a_\mu$ through the similarity of the Feynman diagrams as well as the application of beyond the standard model (BSM) physics to their solutions makes both observables of interest to consider together in this work.

# Chapter 4

# Computational Framework

To achieve the goals of this project the code was structured in such a way that the user first specifies the desired parameters and their range of values which is fed into a section containing the definitions of the observables in the data. The values of these observables calculated from the sampled parameter values are then compared with the measured value from experiment or theory and used to calculate the likelihood for one set of data points. This process is repeated with the marginalized likelihood calculated for each data set and summed over to obtain the result for the whole model. The values of the sampled parameters, the corresponding observable(s) values and Bayes factor are written to a text file which can then be read for the analysis part of the code for plotting.

## 4.1 Inputs for the Models

The aim of this project is to build a framework in which physical models can be input and the Bayes factor be calculated between them. To begin with the mean values of the model parameters are input into a list named *central* with the upper and lower bounds the parameters are to be sampled according to are used for the list labelled $fitparameters$ with the input of the observables being combined under the *observables* list which contains the names of the relevant observables, the observables mean values and the covariance matrix for the observables,

```
central     = {'parameter_1':mean_value_1,
               'parameter_2':mean_value_2,
                 .
                 .
                 .
               'parameter_n':mean_value_n}

fitparameters     = [['parameter_1',lower_bound,upper_bound],
                     ['parameter_2',lower_bound,upper_bound],
                     .
                     .
                     .
                     ['parameter_n',lower_bound,upper_bound]]


observables=  [['Process_Name',
                 ['observable_1_name','observable_2_name,...,
                                              observable_n_name],

                 np.array([observable_1_mean,observable_2_mean,
                                          ...,observable_n_mean]),
                 np.linalg.inv(
                     np.array([[observable_covariance_matrix]]))),
                                               'label']].
```

The parameters are then initialised using the function below which sets the parameters to their mean value to begin the sampling,

```
def __init__(self, central):
    self.para = central
    self.set_parameter([])
```

Where the *set_parameter* function matches the parameter's name to the corresponding sampled value as follows,

```
def set_parameter(self,fitparameters):
    for parlist in fitparameters:
        self.para[parlist[0]] = parlist[1]
    self._set_wc()
```

and the *_set_wc*() takes the corresponding sampled value and converts it so it to a format which can be used in the desired calculations,

```
def _set_wc(self):
        self.fitpar_1=self.para['fitpar_1']
        self.fitpar_2=self.para['fitpar_2']
        .
        .
        .
        self.fitpar_n=self.para['fitpar_n']
```

The effective theory in the code is composed as a list of the parameters. The observable class which contains the calculations reads the parameters as elements of this list. For the case where the parameters are all independent then the effective theory here just contains a list of the sampled parameters however if there is another parameter with dependencies on other sampled parameters then this is input into this list,

```
def write_effective_theory(self, et):
      et.process = [self.fitpar_1,self.fitpar_2,...,
                                              self.fitpar_n]
      return et.process
```

as one can see the parameter list is composed simply of $n$ independt sampled parameters. An example where this is not the case would be,

```
def write_effective_theory(self, et):

et.process = [self.fitpar_1,self.fitpar_2+self.fitpar_3]
return et.process
```

where this contains two parameters but one is the sum of two sampled quantities. To see a case where the parameters are not independently used like this see the test model in chapter 5.

To calculate a theoretical prediction for the observables, an observable class is created containing the equations required for the calculation of the observables.

```
class process_observales:

    def evaluate(self,list,et):
            self.parameter_1=
                    model.write_effective_theory(et)[0]
            self.parameter_2=
                    model.write_effective_theory(et)[1]
            .
            .
            .
            self.parameter_n=
                    model.write_effective_theory(et)[n-1]
            .
            .
            .
            self.observable_1=...
            self.observable_2=...
            .
            .
            .
            self.observable_n=...
```

As mentioned earlier and demonstrated here the parameters used to calculate the observables are defined as elements of the list discussed in the effective theory function. These sampled parameters are then used in the equations input in this class for the observables which then allows one to calculate the likelihood for the model for this set of observables and corresponding parameter points.

## 4.2 Likelihood Calculation

The sampled values are used to calculate the unnormalised log-likelihood as follows,

$$log - likelihood = -\frac{1}{2}(\vec{x} - \vec{\mu}).\Sigma^{-1}.(\vec{x} - \vec{\mu})^T, \tag{4.1}$$

where $\vec{x}$ and $\vec{\mu}$ are row vectors of the calculated observable values and their means respectively with $\Sigma$ representing the covariance matrix (see section 2.5).

The calculation for the log-likelihood starts with examining the list of the fit parameters. When the sample is being run the code performs a check whether the sampled value is within the range specified in the model input. If it is outside of the range, the code simply just returns $-1000000.0$, an arbitrary value returned to simply end the current iteration for an invalid sample set. If it is within the range the code carries on and calculates the log-likelihood.

The Markov-chain Monte-Carlo sampling was performed in our code using the Metropolis-Hastings algorithm which generates points according to a specified target distribution after the chain has been through enough iterations and reached equilibrium. The first step in the algorithm is to define the proposal distribution [25], $Q(s|x)$ from which a candidate, $s$, for the next point in the chain, $x_{t+1}$ can be drawn. Once a candidate sample is selected the code then decides whether to accept or reject the point using the Metropolis ratio [25],

$$\rho = \frac{g(s)Q(x_t|s)}{g(x_t)Q(s|x_t)}, \tag{4.2}$$

where $g$ denotes the target distribution and in our analysis represents the distribution of the observable points after the MCMC sampling has reached equilibrium. Initially the Markov chain produces samples that are not according to the specified target distribution so as such a "burn-in" phase is introduced in which one discards the sampled points that correspond to values produced before the Markov chain has reached equilibrium, usually $\sim \mathcal{O}(1000)$ points are discarded for the burn-in. Note that if parameters are sampled according to a symmetric distribution, such as a Gaussian or uniform distribution, then $Q(x_t|s) = Q(s|x_t)$ which simplifies equation 4.2 so the Metropolis ratio is then given as,

$$\rho = \frac{g(s)}{g(x_t)}. \tag{4.3}$$

The value of $\rho$ determines whether the candidate is accepted or not, if $\rho \geqslant 1$ then it is accepted and we set $x_{t+1} = s$. If $\rho < 1$ then the candidate is only accepted with a probability of $\rho$ by selecting a random number from a uniform distribution $\mathcal{U}(0, 1)$ and comparing it to $\rho$. If a candidate is rejected then we set $x_{t+1} = x_t$ and the process is repeated.

The Markov chain Monte Carlo sample is initiated in Python by first specifying a target, with the proposal and the initial values.

```
log_target = lambda x: unnormalized_log_pdf_gauss(x,
                        fitparameters, et, observables)

mc = pypmc.sampler.markov_chain.AdaptiveMarkovChain(
                                    log_target,prop, start)

mc.run(10**6)
mc.clear()
```

with the number of points to be generated given. Here $10^6$ is used as an example and is typically the order of magnitude used in our analysis.

## 4.3 Saving Output Values

Any output of interest, such as the sampled parameters or observable values are written to a text file which is accomplished in Python by,

```
file=open(``filename.txt,"w")
file.write(example_value_list)
file.close()
```

Here one almost treats the file like a Python package, it is given a name, here for example it is simply named 'file', and then any operations performed with it is initiated by the same notation one would use with any other package. When opening a file in Python it is important to specify what one wishes to do with it. This is done in the second argument of the "open" function, here "w" denotes the intention to write to the file, "r" would mean to read the file and "a" to append it. Next a command is given to the file depending on what the user wishes to do, for example here we are writing a list of values to the text file and then the file is closed as to avoid any unintentional changes. This has the advantage that when the program is finished and closed, the values are stored for analysis again without the need to run through the sampling each time one wishes to perform some analysis with the dataset. This is particularly useful for example when considering large numbers of parameters each with a large number of sampled points as one does not need to keep the large amount of data stored for the entirety of the analysis, one need only read the desired information from the file as desired which saves on computing time and reduces the computing power necessary.

The % operator in Python offers a useful method for saving outputs and calling the relevant ones. It lets one format a string by using the % operator as a variable which saves the need for hard coding separate output names for every possibility, so to write the observable values to a file for a given model, one would use,

```
file=open("observablevalues %s observablevalue.txt"
                                        %(modelname),"w")
str1 = '\n'.join(str(e) for e in obsvallist)
file.write(str1)
file.close()
```

which can be run through a loop for all observables (in this case) and all models instead of hard coding each observable for each model to write to a file separately. In this notation $\%s$ is used to denotes a dynamic string component and $\%(modelname)$ denotes which variable $\%s$ is to be replaced by.

## 4.4 Separate Generating and Analysis Code

The code developed for the project was split into two parts, the code for generating the likelihood, parameters and observable values and the part that was used for analysing these values by plotting or any other analysis tools one wishes to perform. This saves on computer run time by not requiring one to run the full sampling each time one wishes to perform some analysis.

To begin with each model is initialised as follows (we use the MSSM, denoted susy1 as an example). We set the model equal to the one we are interested in. The relevant fitparameters list whose input form was discussed in the general context previously is then called which allows the observables to be calculated after we set the observable list to the corresponding one for the model we are interested in.

```
from susy1data import *

model=susy1(susy1.central)
fitparameters=susy1.fitparameters
observables=susy1.observables
```

Then the physical process we wish to examine is specified,

```
SUSY1process=SUSY1process()
process = {'SUSY1process': SUSY1process}
```

The final part of the generating code is to save the relevant values to files as detailed previously, for example the likelihood, observable values are saved by,

```
likelihoodfile=open("likelihoodvalues %s %s.txt"
                                        %(modelname, z),"w")
str1 = '\n'.join(str(e) for e in likelihood)
likelihoodfile.write(str1)
likelihoodfile.close()

observfile=open("observablevalues %s %s.txt"
                                         %(modelname,z),"w")
str2 = '\n'.join(str(e) for e in observablecalclist)
observfile.write(str2)
observfile.close()
```

The model selection for the analysis part of the code is initiated similarly to that for the generating code except for the omission of the parts required for the parameter sampling,

```
if modelname=='susy1':
    from susy1data import *

    model=susy1(susy1.central)
    fitparameters=susy1.fitparameters
    observables=susy1.observables
```

Since the values required for analysis are saved to text files they need to be accessed by the code for analysis. This is done as follows in the example of the observables which are then saved in the code in individual lists for each set of values,

```
observablecalclist=[]
observfile=open("observablevalues %s %s.txt" %(modelname,z),"r")
for line in observfile:
     l = line.split()
     observablecalclist.append(float(l[0]))
observfile.close()
```

As discussed the analysis section of the code is separated from the code which generates the values. The initial part of the analysis could is outlined below. It starts by importing the code containing the model information relevant for analysis using the << ... >> tool in Emacs which allows one to save chunks of code and copy them in by placing their name in between the angle brackets. A quantity is defined, here $l$ whose only purpose is to fulfil condition on the loops in Python, so while $l$ does not equal some arbitrary value (here we use 10) the code continuously loops around allowing the user to continue performing analyses without requiring to restart the code each time one wishes to analyse something else. The user is then asked to specify the desired model. If the model name that is input is "break" the code is stopped by setting $l = 10$. After the model selection, the code pulls the relevant list of observable names input at the start and displays them to the user who then selects the observable they wish to analyse the parameters of (or indeed the observable itself) whose position on the list of observables is denoted as $t$ below and is calculated in Python merely by using established commands to display the numerical index of the observable in the list. This algorithm is coded as,

```
l=0
while l!=10:
<<ModelsA>>
if modelname=='break':
l=10
else:
obslist=[]
for obs in observables:
    print (obs[1][0], ' : ',obs[4])
obslist.append(obs[1][0])
h=0
while h!=10:
    chosenobservable=raw_input('Observable: ')
if chosenobservable=='break':
    h=10
else:
    t=obslist.index(chosenobservable)
```

After the model the user wishes to analysis is selected, the next step is to select the relevant parameters for the analysis one wish to perform. To do this the

code is structured such that it starts with an empty list *dlist* and then it selects each element, $d$ from the list of fitparameters list whose names were input at the start and append each of these to the list. Finally the list of parameters is printed on screen to allow the user to read off the relevant parameters to their analysis without having to remember which parameters were available from memory,

```
o=observables[t][1][0]
dlist=[]
for d in fitparameters:
      print (d[0])
      dlist.append([d[0]])
```

The axes of the plots in the code are determined by using *raw_input* which allows the user to specify any combination of available components,

```
p1=raw_input('x-axis: ')
parlist1=[]
if [p1] in dlist:
     d1=dlist.index([p1])
for p in parameterlist:
      parlist1.append((p[d1][1]))
```

Above describes how the code allows the user to cycle through the parameters and plot any combination however it is often a desire to plot an observable on one or both of the axes, this is done using *elif* in Python, short for "else if". This tells the code that if the input is not found in the list of the parameters then it should check the list of observables to see if the input matches up with an element from that list. The quantity $t$ below is the same as discussed previously regarding the observable index,

```
elif p1==observables[t][1][0]:
    for o in observablecalclist:
        parlist1.append(o)
```

If the input is still not found in the observable list the code is then instructed to search text files of the form below. This is done to check for the cases where observable values are saved for those observables not in the input list. These cases are those where we consider a model with multiple observables but we wish to consider them individually, for example the case we discuss of the MSSM where the model contains both $a_\mu$ and $Br(\mu \to e\gamma)$. In this

instance the "observablevalues" list would be a vector of size two containing the values for each of these observables. One approach would be to code in a key specifying the $a_\mu$ is the first elements of the vector and $Br(\mu \to e\gamma)$ is the second, however this is not the most general approach as we do not restrict our project to models with only two observables, so for those models where the observables may want to be analysed individually they are saved in individual files and called up using the following code which employs the $\%s$ operator explained previously,

```
else:
    obs1file=open("observablevalues %s %s.txt"
                                        %(modelname,p1),"r")
    for line in obs1file:
        l = line.split()
        parlist1.append(float(l[0]))
    obs1file.close()
```

This algorithm laid out above allows one to plot any combination of parameters, observables or even observables against parameters without the need to hard code the combination each time the user wishes to generate a different combination of axes.

## 4.5 Plotting

The mathplotlib package in Python offers numerous plotting options. Whilst writing the program one of the main objectives was to make the code adaptable and limit the hard coding needed when inputting a new model, this goal was achieved for the plotting section of the code by utilising the % operator discussed previously.

The analysis code starts by printing the list of all models input into the code with a variable called *modelname* defined as the choice the user selects from the list,

```
modellist=[...]
    for m in modellist:
    print (m)
modelname=raw_input("Model: ")
```

A similar approach is taken when the user decides which elements of the model to plot, a list of parameters of the specified model is printed and two are selected to form the axes of the plots. This allows the use of the same code for plotting regardless of which model and observables are being considered. Similarly to how the parameter and other values are written to files by using the % operator, the plots are saved by Python using this method which saves the need to hard code what is on the axes and which model into the file name each time,

```
plt.savefig('%s vs %s %s %s.png' %(p1, p2, modelname, o))
```

Here $p1$ and $p2$ are the variables being plotted, so for example if a plot was generated of $a_\mu$ and $Br(\mu \to e\gamma)$ in the MSSM then this would save the graph as '$a_\mu$ vs $Br(\mu \to e\gamma)$ MSSM.png'.

The % operator has been shown how it can be useful when one requires a string in Python to be changeable, however some aspects of the code did not require this. For example the labels on the axes of the plots did not need to be hard coded for each possible outcome but it also did not require the % operator. When decided what to plot the user is presented with the list of parameters and observables of the model as discussed earlier with the selection being made by using the *raw_input*('...') command which prints the argument on screen to indicate to the user what to input. It then saves whatever the user has typed as a response as a string. This string is then matched to the relevant quantity in the code whose values are then used for the axis of the plots,

```
p1=raw_input(...)
p2=raw_input(...)
plt.xlabel(p1, fontsize=...)
plt.ylabel(p2, fontsize=...)
```

Hex plots were one of the plotting types used in our analysis as they are useful to represent the relationship of two variables, one of the advantages over standard scatter plots is that the bins are shaded depending of how many points lie in the bin rather than just having a sea of overlapping points in a scatter plot. They also have the advantage of each bin being correlated to the un-normalised log-likelihood of the model thereby allowing one to easily see the most likely values of the quantities on the axes by a quick inspection of the graphs to see where the most shaded areas lie (see the output produced

in chapter 5 for examples of hex plots). The hex plots were constructed using the code of the form,

```
x=parlist1
y=parlist2
plt.hexbin(x, y, gridsize = 200)
cb = plt.colorbar()
plt.xlabel()
plt.ylabel()
cb.set_label('counts')
plt.xlim([])
plt.ylim([])
plt.show()
```

Scatter plots were the other main type of plots produced. These were used for plots where the likelihood of the quantities was not of interest (see Fig 6.2 in section 6.4 for an example of this). Scatter plots are constructed similarly to hexbin plots however *plt.hexbin*(...) is replaced with *plt.scatter*(...) with no need for a gridsize or colour bar to be specified,

```
x=parlist1
y=parlist2
plt.scatter(x,y)
plt.xlim([])
plt.ylim([])
plt.xlabel()
plt.ylabel()
plt.show()
```

# Chapter 5

# Test Model

An important task was to first test the initial parts of the framework. To do this, the data from [26] containing plots of observables relating to $b-$decays was used, this was chosen because of the possible indicators of deviations in the quark sector for experimental Standard Model predictions [27, 28]. The observables used were from $b$ decays with Hamiltonian ,

$$\mathcal{H}_{eff}^{\bar{1}u_2d_1} = \frac{G_F}{\sqrt{2}} V_{u_1b} V_{u_2d_1}^* [C_1 Q_1^{\bar{u}_1 u_2 d_1} + C_2 Q_2^{\bar{u}_1 u_2 d_1}] + ... \tag{5.1}$$

where $G_F$ is the Fermi constant and the ellipses denote penguin operators not considered here leaving two tree level operators, the colour singlet operator $Q_1$ and the colour rearranged operator $Q_2$ which have the form,

$$Q_1^{\bar{u}_1 u_2 d_1} = (\bar{u}_1^\alpha b^\beta)_{V-A} (\bar{d}_1^\beta u_2^\alpha)_{V-A}, \tag{5.2}$$

$$Q_2^{\bar{u}_1 u_2 d_1} = (\bar{u}_1^\alpha b^\alpha)_{V-A} (\bar{d}_1^\beta u_2^\beta)_{V-A}. \tag{5.3}$$

where $\alpha$ and $\beta$ are colour indexes and the $V-A$ component has been abbreviated as $(\bar{q}q') = \bar{q}\gamma_\mu(1-\gamma_5)q'$ with $\gamma$ denoting the usual gamma matrices. Since this project is concerned with comparing models in a statistically meaningful manner according to experimental data it made sense to test the components of the code with a model composed of previously analysed observables that can be parameterised by some new physics contributions. The tree level Wilson coefficients, $C_1$ and $C_2$ given in equation (5.1) can be extended from their Standard Model values by the addition of a new physics contribution,

$$C_1 = C_1^{SM} + \Delta C_1, \qquad C_2 = C_2^{SM} + \Delta C_2, \tag{5.4}$$

where $C_1^{SM}(\bar{m}_b) = -0.31$, $C_2^{SM}(\bar{m}_b) = 1.1$ [40] and $\Delta C_i$ denotes the new physics contribution. The effects of these Wilson coefficients having new physics extensions is largely constrained by observables related to different decays of $b-$quarks.

The aim of this chapter is to examine experimental results for numerous observables depending on the Wilson coefficients $C_1$ and $C_2$ and examine the effect of the new physics contribution to these coefficients. Plots of these observables are created showing how they constrain the real and imaginary parts of the new physics contributions, $Re(\Delta C_1)$ and $Im(\Delta C_1)$, to the Wilson coefficient $C_1$. The constraints on the values of $Re(\Delta C_1)$ and $Im(\Delta C_1)$ is then examined when all the observables are combined showing the region of values corresponding to those of highest probability in the model as a whole.

## 5.1 $b \to uud$

The decay width of the $b-$quark is given by the expression from Ref. [28] in which the explicit dependence on the Wilson coefficients $C_1$ and $C_2$ can be seen,

$$\Gamma = \frac{3|C_1|^2 + 3|C_2|^2 + 2Re[C_1^* C_2]}{3|C_1^{SM}|^2 + 3|C_2^{SM}|^2 + 2Re[C_1^{SM*} C_2^{SM}]} \times \Gamma^{SM}, \tag{5.5}$$

where $\Gamma^{SM}$ is the value of the decay width in the Standard Model and the value $(3.6 \pm 0.8) \times 10^{-13}$ GeV [87] and the experimentally measured value for the decay width, denoted $\Gamma_{tot}$ is taken as $(4.2 \pm 0.02) \times 10^{-13}$ GeV [29]. This allows one to see the effect of introducing new physics contributions to the Wilson coefficients and comparing this to experimental measurements of the decay $b \to uud$.

Additional bounds on the $B-$meson observables can be obtained by introducing the semi-leptonic asymmetry which is calculated as [26],

$$a_{sl}^d = Im\left(\frac{\Gamma_{12}^d / M_{12}^d}{\Gamma_{12}^{d,SM} / M_{12}^{d,SM}} \cdot \frac{\Gamma_{12}^{d,SM}}{M_{12}^{d,SM}}\right). \tag{5.6}$$

References [88, 89, 90] produced results which were used in [26] to give an expression used to calculate the new physics contribution to equation 5.6,

$$\begin{aligned}\frac{\Gamma_{12}^d/M_{12}^d}{\Gamma_{12}^{d,SM}/M_{12}^{d,SM}} = 1 &- (0.23 - 0.047i).\Delta C_1 + (0.76 + 0.25i).\Delta C_1^2 \\ &+ (1.91 - 0.0029i).\Delta C_2 + (0.084 + 0.14i).\Delta C_1.\Delta C_2 \\ &+ (0.93 + 0.0072i).\Delta C_2^2.\end{aligned} \tag{5.7}$$

The Standard Model value for $\Gamma_{12}^q/M_{12}^q$ was calculated in [30] to be,

$$\frac{\Gamma_{12}^{d,SM}}{M_{12}^{d,SM}} = -0.0050 - 0.00045i. \tag{5.8}$$

The fact that equation (5.8) is complex is a consequence of the fact that $B^q - \bar{B}^q$ is parameterised by diagonalising $M - i\Gamma$ where $M$ is the mass matrix and $\Gamma$ is the decay matrix. This results in the argument of equation (5.6) being complex **??**. The mean value for semi-leptonic asymmetry taken to be the averaged results calculated in references [31, 32, 33, 34, 35, 36],

$$a_{sl}^d = (2.2 \pm 2.2) \times 10^{-3}. \tag{5.9}$$

which can then be used in our analysis when calculating the likelihood of this observable, the method of which is discussed later in section 5.7.

## 5.2 $B \to \pi\pi$

Another decay examined in this model was $B \to \pi\pi$ which can be calculated as a tree level decay. A constraint on this decay can be found in the following ratio provided by [37] (using [38]) ignoring small contributions from electroweak penguin diagrams,

$$R_{\pi^-\pi^0} = \frac{\Gamma(B^- \to \pi^-\pi^0)}{d\Gamma(\bar{B}^0 \to \pi^+\ell^-\bar{\nu}_\ell)/dq^2|_{q^2=0}}$$

$$\simeq (0.70^{+0.12}_{-0.08})[1 + 2.20Re(\Delta C_+^{uu}) - 0.13Im(\Delta C_+^{uu}) + 1.21|\Delta C_+^{uu}|^2]GeV^2, \tag{5.10}$$

where $\Delta C^{pp} \equiv \Delta C_1^{pp} + \Delta C_2^{pp}$. Here $\Delta C_i^{pp}$ denotes the new physics contribution to the Wilson coefficient $C_i^{pp}$. In the Standard Model the Wilson coefficients $C_{1,2}^{pp}$ do not depend on the quark flavour, however it is possible that new physics contributions to these coefficients do depend on the quark flavour. Since this

section was concerned with testing the sampling and plotting aspects of the code we follow the assumptions laid out in [26] mainly that we consider a model where $\Delta C_{1,2}$ are flavour independent but for completeness we include the flavour indices in the relevant equations. The ratio has an experimentally measured value of $R_{\pi^-\pi^0} = (0.81 \pm 0.14)\text{GeV}^2$ [37] which is used as the observables mean value in our analysis.

## 5.3 $B \to D^{(*)0}h^0$

Since this one of the aims of this chapter is to reproduce the results of Ref [26] to serve as a check for our computer framework the observable relating to the decay $B \to D^{(*)0}h^0$ (where $h^0 = \{\pi^0, k^0\}$) is also considered. This decay allows for the analysis of another observable which constrains the Wilson coefficients, the indirect CP asymmetry [40],

$$S_{D^{(*)0}h^0} = 2\frac{Im(e^{-2i\beta}\rho_{D^{(*)0}h^0})}{1+|\rho_{D^{(*)0}h^0}|^2}, \qquad \rho_{D^{(*)0}h^0} = \frac{C_1^{uc} + C_2^{uc}/3}{(C_1^{uc})^* + (C_2^{uc})^*/3}. \quad (5.11)$$

Where in our analysis the flavour indices on the Wilson coefficients are not relevant since we consider a new physics model where the $\Delta C_i$ is flavour independent however for completeness we leave the indices in. The experimental measurement for this indirect asymmetry is given as [39],

$$S_{D^{(*)0}h^0} = -(56 \pm 24)\% \quad (5.12)$$

which acts as the mean value and standard deviation in our construction of the likelihood function for this observable.

## 5.4 $\bar{B}^0 \to D^{*+}\pi^-$

Another observable considered in Ref. [26] was obtained from examining the decay $\bar{B}^0 \to D^{*+}\pi^-$. The observable associated with this decay is the ratio of the decay rate of $\bar{B}^0 \to D^{*+}\pi^-$ to the differential semi-leptonic $\bar{B}^0 \to D^{*+}l^-\nu^-$rate

$$
\begin{aligned}
R_{D^{*+}\pi^-} &= \frac{\Gamma(\bar{B}^0 \to D^{*+}\pi^-)}{d\Gamma(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell)/dq^2|_{q^2=m^2_{\pi^-}}} \\
&\simeq (1.07 \pm 0.04)[1 + 1.90Re(\Delta C^{uc}_{1/3}) + 0.02Im(\Delta C^{uc}_{1/3}) \\
&\quad + 0.09|\Delta C^{uc}_{1/3}|^2]\ \text{GeV}^2,
\end{aligned}
\tag{5.13}
$$

where the experimental value for the ratio is $R_{D^{*+}\pi^-} = (0.96 \pm 0.08)GeV^2$ and $\Delta C^{pp'}_{1/3} = \frac{\Delta C^{pp'}_1}{3} + \Delta C^{pp'}_2$.

## 5.5 $B \to \rho\rho$

The last observable used in this test model was the ratio of the branching ratio $B \to \rho^-\rho^0$ to $B \to \rho^+\rho^0$ and was used as the final constraint of the parameter space in Ref. [26]. The dependence on the Wilson coefficients can be seen by examining the expression for this observable taken from Ref. [29],

$$
\begin{aligned}
&R(\rho^-\rho^0/\rho^+\rho^-) = \\
&\qquad (0.65^{+0.16}_{-0.11})\left[\frac{1 + 2.1Re(\Delta C^{uu}_+) + 0.06Im(C^{uu}_+) + 1.1|C^{uu}_+|^2}{1 + 2.0Re(\Delta C^{uu}_{1/3}) + 0.23Im(\Delta C^{uu}_{1/3}) + 1.0|\Delta C^{uu}_{1/3}|^2}\right].
\end{aligned}
\tag{5.14}
$$

which has an mean value and standard deviation of [40],

$$
R(\rho^-\rho^0/\rho^+\rho^-) = 0.89 \pm 0.14. \tag{5.15}
$$

which then allows, with the combination of all the other observable mean values and standard deviations, the construction of a likelihood function for this model as a whole allowing one to plot the overlap region of the new physics contributions to the Wilson coefficients.

## 5.6 Scaling of the Wilson Coefficients

The Wilson coefficients $C_{1,2}$ are at the scale of the bottom mass in the calculations of the observables above, however these coefficients are obtained at the scale of the $W$ boson scale by matching to those at the new-physics scale through renormalization-group evolution [40]. For ease bounds placed on these Wilson coefficients was done at the $M_W$ scale, this required an expression that

could be used in our analysis to make the scale of $C_i$ interchangeable [40],

$$\Delta C_1(m_b) = z_+ \Delta C_1(M_W) + z_- \Delta C_2(M_W), \tag{5.16}$$

$$\Delta C_2(m_b) = z_- \Delta C_1(M_W) + z_+ \Delta C_2(M_W), \tag{5.17}$$

where,

$$z\pm = \frac{1}{2}(\eta_5^{\frac{6}{23}} \pm \eta_5^{\frac{-12}{23}}), \quad \eta_5 \equiv \frac{\alpha_s(M_W)}{\alpha_s(m_b)}. \tag{5.18}$$

Running the sample at the scale $m_b$ for the Wilson coefficients and taking $\Delta C_2(M_W) = 0$ gives,

$$\Delta C_2(m_b) = z_- \Delta C_1(M_W), \tag{5.19}$$

from the expressions above which allows us to generate the observables with the same constraints as in [26]. When looking at bounds on $\Delta C_1(M_W)$, $\Delta C_2(M_W)$ is set to zero to allow their impact on the observables to be explicitly viewed.

## 5.7 Model Results

The paper [26] presented the plot which we used to compare results. Each of the coloured regions corresponds to a different observable constrained by the Wilson coefficients the individual plots from our analysis are given in Fig. 5.2. The formula and data for these was used separately first to check all the regions produced the correct result in our analysis and then they were combined to show the overlap which gives the overall allowed region of the Wilson coefficients shown in yellow in Fig. 5.1. To calculate the overlap region shown in Fig. 5.1 the observables and their experimental values and errors were input into a list and covariance matrix respectively.

To begin with the observables were considered individually to allow one to see the individual coloured regions in Fig. 5.1. For each observable a Gaussian distribution is used as the fitness function in the Markov-chain, $g$ in equation 4.2. Recall equation 2.7 for a one dimensional Gaussian,

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

When considering each observable to produce the plots in Fig. 5.2 the mean value was used for $\mu$ and the error used for $\sigma$ so for the example of the decay

width of $b \to uud$, $\mu = 4.2 \times 10^{-13}$ GeV and $\sigma = 0.02 \times 10^{-13}$ GeV with $x$ being the calculated value in our code from the sampled parameters. The results from these observables are shown in Fig. 5.4 with different colours corresponding to the observables and the results from our analysis are shown in Fig. 5.2.

To plot the overlap region of all the observables a multi-dimensional Gaussian is used as the target distribution. Equation 2.9 gives the following expression for an n-dimensional Gaussian,

$$P(\vec{x}) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} e^{\left(-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T\right)},$$

To analyse the overlap region a vector of the mean values of the observables is constructed,

$$\vec{\mu} = [4.2 \times 10^{-13}, 2.2 \times 10^{-3}, 0.81, 0.96, -0.56, 0.89] \tag{5.20}$$

which represent each observables mean value as discussed earlier in this chapter. Since the observables are uncorrelated then from equation 2.8 the covariance matrix is going to be a diagonal matrix. The fact that the covariance matrix is diagonal means that it can be constructed by simply inputting the experimental errors as the diagonal elements like so,

$$\Sigma = \begin{bmatrix} (0.02 \times 10^{-13})^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & (2.2 \times 10^{-3})^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.14^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.08^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.24^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.14^2 \end{bmatrix}$$

where $\Sigma_{ii}$ corresponds to the experimental error of $\mu_i$ and are taken from the relevant sections previously. This covariance matrix along with the vector of mean values allows us then to construct a Gaussian target distribution for use in the MCMC sampling represented by the function $g()$ in equation 4.2. The sample ranges for $\Delta C_i(M_W)$ is taken to be $[-2, 2]$ so to more than cover the plot ranges in figures 5.1 and 5.4. Note this range is extended for Fig. 5.3 to be the ranges on the axis. This then produces the overlap plot outlined in yellow in Fig. 5.1 and recreated in our analysis in Fig. 5.4.

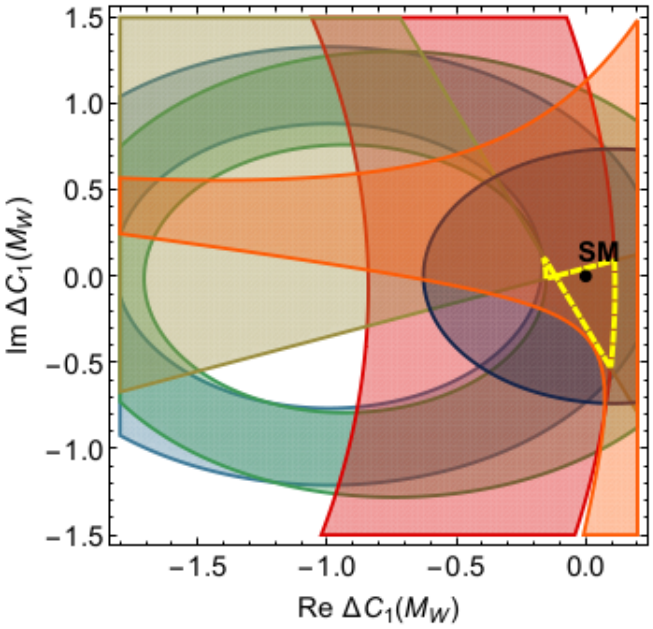



**Figure 5.1:** *Results from [26] demonstrating the constraints on the parameters from the relevant observables. This plot consists of individual observable constraints overlaid and the overlap region outlined in yellow. These individual observable constraints were recreated and shown previously in figures 5.1-5.5.*

Figure 5.2 shows the values of $Re(\Delta C_1)$ and $Im(\Delta C_1)$ constrained by the mentioned observable. One can see that the observables favour deviations from the Standard Model result $Re(\Delta C_1) = Im(\Delta C_1) = 0$. Figures 5.2b, 5.2d and 5.2e seem to show an incomplete picture of the behaviour of these observables but appear to allow for even greater deviations from the Standard Model than the other observables. Figure 5.3 shows these three plots over a greater range to try and give a more complete picture specifically how much more of a deviation from the Standard Model is allowed under these observables. One can also see the full shape of Fig. 5.2f when plotted over a greater range as shown in Fig. 5.3c and the resemblance of figures 5.2f and 5.2c. Figures 5.2c 5.2e and 5.2f are in the shape of a ring with values of $Re(\Delta C_1)$ and $Im(\Delta C_1)$ inside the ring corresponding to the relevant observables values being far smaller than the mean and values outside the ring being far larger, hence the ring of preferred values. The different shape in Fig. 5.2b is better explained by examining the larger picture given by Fig. 5.3a with values lying in the white upper left and lower left quadrants corresponding to values of $a_{sl}^d$ being far larger than the mean value and values in the right two quadrants corresponding to values

being far lower than the mean. For Fig. 5.2d both the upper and lower blank regions correspond to values far greater than the mean value for $S_{D^{(*)0}h^0}$ from equation (5.12) with the straight borders on the plot a result from the ratio of the Wilson coefficients from $\rho_{D^{(*)0}h^0}$ in equation (5.11).

Fig. 5.4 shows the overlap region of the observables outlined previously. We found that our sampling method in Python produced the same overlap region as that of [26] given the same input as expected. This served the purpose of checking the sampling and plotting aspects of the code.

It was prudent to test the statistical part of the code while the observables and data were already coded in. To do this we took the input from the observables above and changed how the parameters were defined, so where as previously we had $\Delta C_i = Re(\Delta C_i) + iIm(\Delta C_i)$ for the Wilson coefficients, this was changed to something deliberately incorrect $\Delta C_2 = Re(\Delta C_2) + iIm(\Delta C_2) + iIm(\Delta C_1)$. Using this erroneous definition gave the overlap region for the observables given in Fig. 5.5.

The produced results were used to calculate the Bayes factor between the model defined above composed of the $B-$physics observables and the so called "wrong-model" also composed of the same observables but with the changed definition of the Wilson coefficients. The numerical output for this was $\mathcal{O}(10^{20})$ against this "wrong-model". A value $> 100$ was to be expected given the fact this model was deliberately constructed to produce incorrect results and from the details in the table of interpretation for the Bayes factor, a value of this magnitude indicates that the original model containing the $B-$physics observables is decisively more preferred given the data.

After examining this model and the constraints on the observables one may then be interested in how this compares to the Standard Model. As discussed previously, the observables shown in Fig. 5.2 that constrain $\Delta C_1$ show significant allowed deviations form the Standard Model value of $\Delta C_1$. By inspection of these graphs alone and especially Fig. 5.4 one can see the significant deviation from the Standard Model value and could conclude that this model of new physics would be favoured over the Standard Model however this can be shown more rigorously. Since the observables have explicit dependencies on $\Delta C_1$ and $\Delta C_2$ then for most of the observables this results in a calculated

value of zero. The form of the Gaussian likelihood is given in equation 2.7 and one can see that when the calculated value is zero this becomes,

$$P(D|x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(-\mu)^2}{2\sigma^2}}. \tag{5.21}$$

For the example of $R_{D^{*+}\pi^-}$ the expression for which is given in equation 5.13 one can see the resulting calculation is zero. Taking $\mu = 0.96$ GeV$^2$ and $\sigma = 0.08$ GeV$^2$ gives a resulting likelihood in the Standard Model of

$$P(D|x) = \frac{1}{\sqrt{2\pi \times 0.08^2}} e^{-\frac{(-0.96)^2}{2\times 0.08^2}} = 7.20 \times 10^{-62}. \tag{5.22}$$

From this one can see that by either using a total expression for the model likelihood like that of equation 2.9 or by multiplying the uncorrelated probabilities (both are equivalent and obtain the same result) one can see that the total likelihood for the Standard Model with these observables is,

$$P^{\text{SM}}(D|\vec{x}_{\text{SM}}) \to 0. \tag{5.23}$$

Recall the expression of the Bayes factor used in our analysis,

$$B_{12} = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\sum_{\vec{\theta}_1} P(D|\vec{\theta}_1) P(\vec{\theta}_1|M_1)}{\sum_{\vec{\theta}_2} P(D|\vec{\theta}_2) P(\vec{\theta}_2|M_2)}.$$

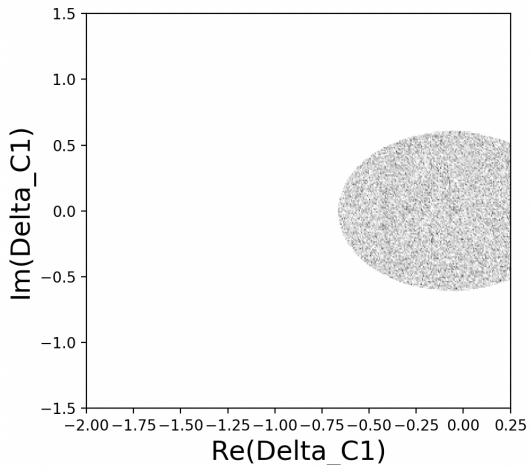
So for the case of $P^{\text{SM}}(D|\vec{x}_{\text{SM}}) \to 0$ one can see this leads to a value of the Bayes factor of,

$$B_{\text{SM,Test Model}} = 0. \tag{5.24}$$

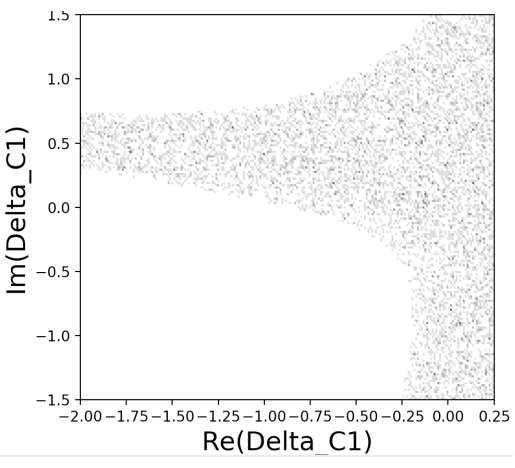
Using table 2.1 this indicates decisively favourability against the Standard Model when considering these observables. This result is to be expected as discussed earlier from inspection of the plots showing the favoured regions of the new physics contributions to $C_1$.

This chapter has demonstrated all the key components of this project, mainly that of sampling, calculations in python, plotting, and finally the Bayes factor. The calculations of each of the observables was a good starting point in the analysis as the plots for how these constrain the new physics contributions to $C_1$ served as a good check to start since the data and plots for these were available in the literature making it a good place to build a foundation on

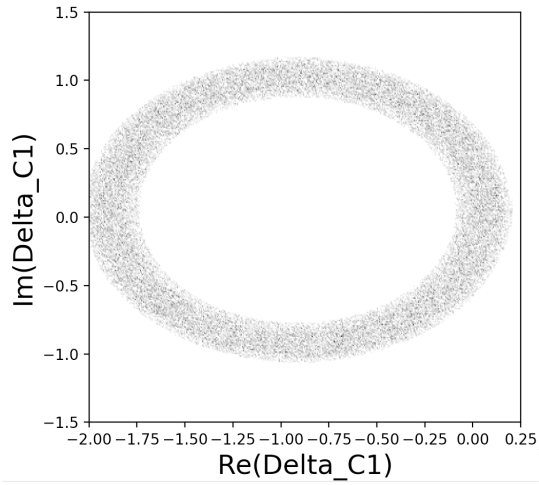
so these tools could be applied in other areas of analyses. The method of constructed a Gaussian target distribution was shown and how this can be achieved when one wishes to deal with multiple observables at once and the role this plays in the calculation of the Bayes factor is demonstrated with this chapter ended with the calculation of the Bayes factor between the Standard Model and this test model. One can see a value of the Bayes factor in this area is to be expected by inspecting the plots and observing significant allowances for deviations from the Standard Model values of $Re(\Delta C_1)$ and $Im(\Delta C_1)$.



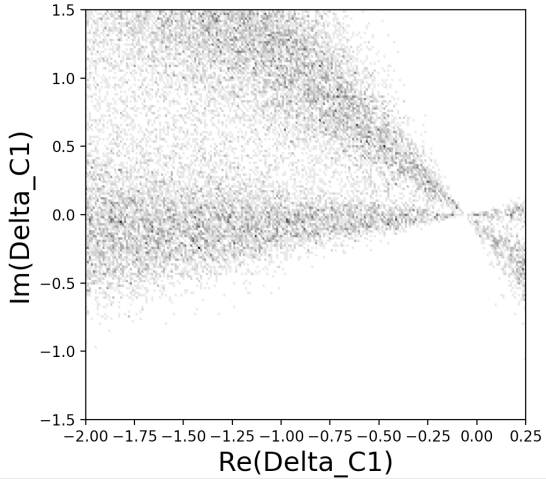

**(a)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the decay width $\Gamma$ given in equation (5.5).*



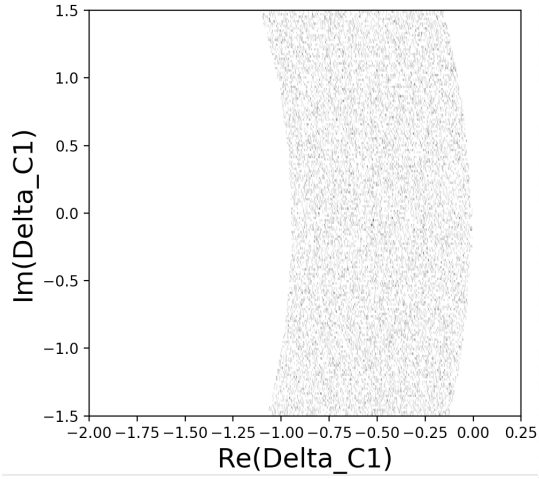

**(b)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the semileptonic asymmetry $a_{sl}^d$ given in equation (5.6).*



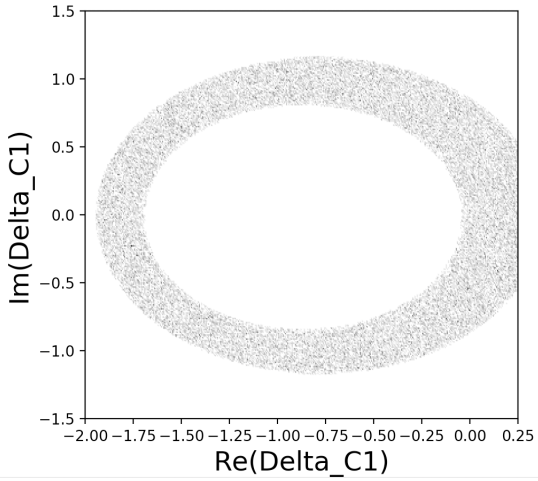

**(c)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the ratio $R_{\pi^-\pi^0}$ given in equation (5.10).*




**(d)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by observable related to indirect CP asymmetry denoted $S_{D^{(*)}h^0}$ for the decay $B \to D^{(*)}h^0$ given in equation (5.11).*



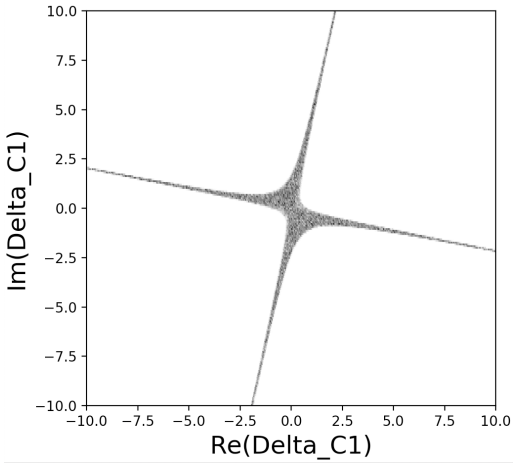

**(e)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the ratio $R_{D^{*+}\pi^-}$ given in equation 5.13.*



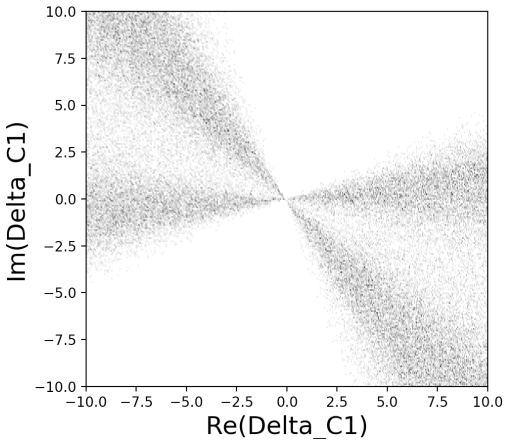

**(f)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the ratio $R(\rho^-\rho^0/\rho^+\rho^-)$ given in equation (5.14).*



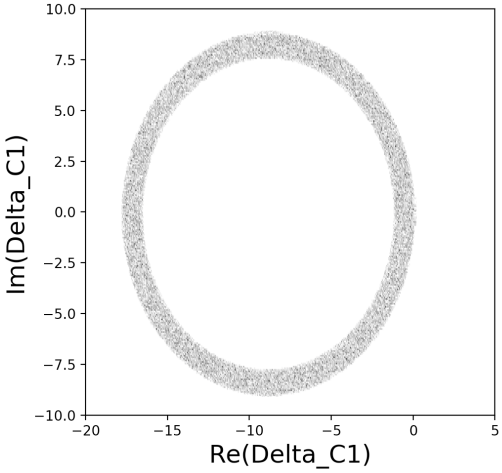
**Figure 5.2:** *Plots showing constraints placed on $Re(\Delta C_1)$ and $Im(\Delta C_1)$ from individual observables.*




**(a)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the semileptonic asymmetry $a_{sl}^d$ given in equation 5.6 and plotted over a larger range than Fig. 5.1.*




**(b)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by observable related to indirect CP asymmetry denoted $S_{D^{(*)}h^0}$ for the decay $B \to D^{(*)}h^0$ given in equation 5.11 and plotted over a larger range than Fig. 5.1.*




**(c)** *Plot showing $Im(\Delta C_1)$ against $Re(\Delta C_1)$ constrained by the ratio $R(\rho^-\rho^0/\rho^+\rho^-)$ given in equation 5.14 and plotted over a larger range than Fig. 5.1.*

**Figure 5.3:** *Plots showing constraints on $Re(\Delta C_1)$ and $Im(\Delta C_1)$ on observables over larger area.*

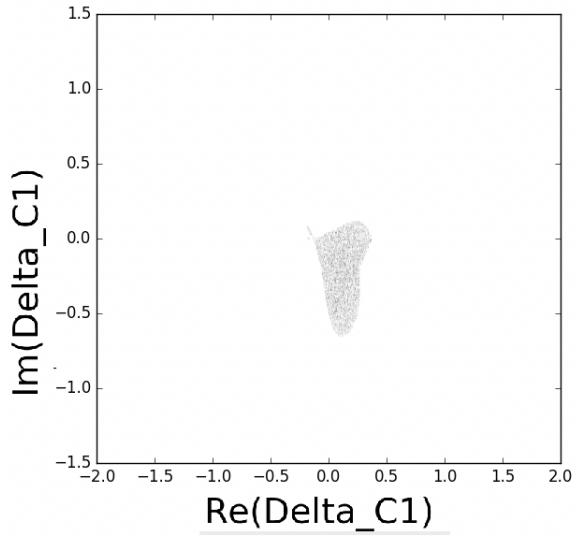



**Figure 5.4:** *Hexbin plot of $Im(\Delta C_1)$ against $Re(\Delta C_1)$ for all observables, c.f. region outlined in yellow on figure 5.1.*

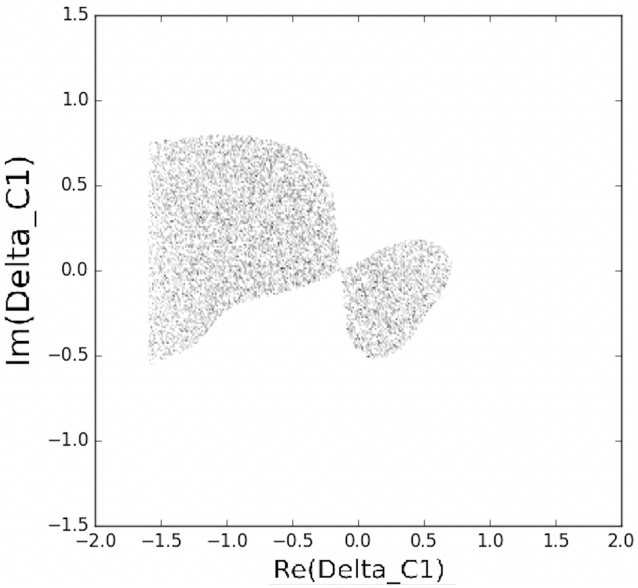



**Figure 5.5:** *Plot of $Im(\Delta C_1)$ against $Re(\Delta C_1)$ for dummy model.*

# Chapter 6

# Minimally Supersymmetric Standard Model

When selecting models to compare for the Bayes factor it was important to select two models which are studied extensively and are therefore able to explain the same experimental data so a comparison between the models can be formed. The two models chosen were the minimally supersymmetric Standard Model (MSSM) and the two-Higgs doublet model (2HDM) since they offer possible explanations for some of the same deficiencies of the Standard Model. The MSSM and 2HDM also represented two of the simplest extensions to the Standard Model making them prime candidates for initial comparisons of this kind with the possibility of comparing more complex models in future work.

The MSSM is the extension to the Standard Model achieved by incorporating $\mathcal{N} = 1$ softly broken global supersymmetry which involves an extension to the Higgs sector to include two Higgs doublets, denoted by $H_u$ and $H_d$ with hypercharge $1/2$ and $-1/2$ respectively [41]. To achieve this extension one introduces a superparticle partner for the particles of the Standard Model whose spin differs by one half and their masses by some order $\mathcal{O}(M_{SUSY})$ whilst keeping all other quantum numbers the same [42]. The superpartners to the matter fields are aptly called the sfermions made up of sleptons and sfermions which are scalar particles with spin zero and the superpartners of the gauge bosons, called the gauginos, are similarly chose to have spin one half. To undertake analysis of this kind it is important to first specify which parameters were used in the calculations. For the initial analysis in the MSSM the parameters

sampled over were,

$$M_1,\ M_2,\ \mu,\ m_{\tilde{L}_{11}},\ m_{\tilde{L}_{22}},\ m_{\tilde{R}_{11}},\ m_{\tilde{R}_{22}}\ ,$$

denoting the bino mass, wino mass, Higgsino mass, and the flavour diagonal soft mass parameters respectively. The use of these parameters to calculate the relevant observables is discussed in throughout this chapter, beginning in the next section with the anomalous magnetic moment of the muon, $a_\mu$ and how this is decomposed in the MSSM into components involving the charginos and the neutralinos. The forms of these constituents of $a_\mu$ are given explicitly along with the relevant loop integrals and the dependencies on other sparticle masses which involve the diagonalising of the relevant mass matrices.

Lepton flavour violating observables offer possible indications into physics beyond the Standard Model. In the MSSM multiple LFV decays have non-negligible branching ratios. As mentioned earlier the similarity in the Feynman diagrams between $a_\mu$ and $Br(\mu \to e\gamma)$ make them intriguing observables to consider together especially in the MSSM where one can parametrise part of the calculation of $Br(\mu \to e\gamma)$ as involving a component of $a_\mu$ as shown in equation (6.39). In addition the experimental limit on $Br(\mu \to e\gamma)$ constrains the flavour violating mass $m_{\tilde{L}_{12}}$ used also in the calculation of $a_\mu$. These links between the two observables make them an interesting topic of joint study. In addition to $\mu \to e\gamma$ other LFV observables are of interest here. In the search for a model to explain the deficiencies of the Standard Model the consideration of $\mu - \tau$ flavour violation is motivated by the observed atmospheric neutrino oscillations $\nu_\mu \leftrightarrow \nu_\tau$. The decay $Br(\tau \to \mu\gamma)$ can be of interest to study with $\mu \to e\gamma$ because of the similarity in the Feynman diagrams with the flavour masses being one generation higher, $m_\mu \to m_\tau$ and $m_e \to m_\mu$, however $Br(\tau \to \mu ee)$ and $Br(\tau \to 3\mu)$ are also of interest to study their effect on the MSSM parameter space and the impact this has on $a_\mu$ and the other LFV observables. The large overlap of the parameter space between the different LFV decays and $a_\mu$ in the MSSM makes these observables intriguing to study together. The impact of examining experimental constraints on these different observables and the effect each of these have on restricting a largely overlapped parameter space is a good exercise for one to perform as a precursor to calculating the Bayes factor as it offers an indication as to which observables constrain the model more and which have a greater impact on the Bayes factor.

As a test of the project so far the results of a plot from [23] showing $Br(\mu \to e\gamma)$ against $a_\mu$ is reproduced to demonstrate the capability of the project to produce consistent results with literature following the same assumptions. This serves as a check of the sampling aspect of the code as well as ensuring the calculation for the observables is being performed correctly. After this constrained plot is reproduced, the effect of loosening these restrictions and the introduction of additional LFV observables is examined on the parameter space..

## 6.1 $a_\mu$ in the MSSM

The anomalous magnetic moment of the muon is a flavour conserving observable which could be explained by the introduction of light sleptons, charginos and neutralinos [23]. The full one loop result was calculated in Ref. [43] and can be broken down into the chargino and neutralino contributions,

$$a_\mu = \sum_k a_\mu^{\tilde{\chi}_k^\pm} + \sum_i a_\mu^{\tilde{\chi}_i^0}. \tag{6.1}$$

This decomposition represents splitting the analytic expressions into those calculated from the neutralino-smuon and the chargino-sneutrino diagrams which allows one to input the formula for $a_\mu$ into the code with more ease.

### 6.1.1 Chargino Contribution

The chargino contribution can be conveniently broken down into two terms as,

$$a_\mu^{\tilde{\chi}_k^\pm} = a_\mu^{\tilde{\chi}_k^\pm}{}_1 + a_\mu^{\tilde{\chi}_k^\pm}{}_2 \tag{6.2}$$

where,

$$a_\mu^{\tilde{\chi}_k^\pm}{}_1 = \frac{m_\mu^2}{192\pi^2 m^2_{\tilde{\chi}_k^\pm}}(g_2^2|V_{k1}|^2 + Y_\mu^2|U_{k2}|^2)[\sin^2(\theta_{\tilde{\nu}})x_{k1}F_1^C(x_{k1}) + \cos^2(\theta_{\tilde{\nu}})x_{k2}F_1^C(x_{k2})], \tag{6.3}$$

$$a_{\mu}^{\tilde{\chi}_k^{\pm}}{}_2 = -\frac{2m_\mu}{48\pi^2 m^2_{\tilde{\chi}_k^{\pm}}} g_2 Y_\mu Re[V_{k1}U_{k2}][\sin^2(\theta_{\tilde{\nu}})x_{k1}F_2^C(x_{k1}) + \cos^2(\theta_{\tilde{\nu}})x_{k2}F_2^C(x_{k2})]. \tag{6.4}$$

The matrices $U$ and $V$ are eigenvectors of the chargino mass matrix which are given in equation 6.14 and the angle $\theta_\nu$ given by equation 6.11. The gauge and Yukawa couplings, $g_2$ and $Y_\mu$ used above in equations 6.3 and 6.4 are given by [45],

$$g_2 = \frac{e}{\sin(\theta_w)} \tag{6.5}$$

$$Y_\mu = \frac{m_\mu g_2}{\sqrt{2}M_W \cos\beta}. \tag{6.6}$$

where $e$ is the elementary electron charge and $\theta_w$ is the Weinberg angle. An important point to note is that in our analysis we worked in natural units, $\bar{h} = c = \epsilon_0 = 1$ which gives a value,

$$e = \sqrt{4\pi\alpha_{em}\bar{h}c\epsilon_0} = \sqrt{4\pi\alpha_{em}} \simeq 0.30282, \tag{6.7}$$

and the mass ratio $x_{ki}$ represents the squared ratio of the chargino to sneutrino mass,

$$x_{ki} = \frac{m^2_{\tilde{x}_k^{\pm}}}{m^2_{\tilde{\nu}_i}}. \tag{6.8}$$

The loop functions used for the chargino components, $F_1^C$ and $F_2^C$, are defined as,

$$F_1^C(x) = \frac{2}{(1-x)^4}[2 + 3x - 6x^2 + x^3 + 6x\ln(x)], \tag{6.9}$$

$$F_2^C(x) = \frac{3}{2(1-x)^3}[3 - 4x + x^2 + 2\ln(x)].$$

The mass ratios used in equation (6.3) and defined in (6.8) involve both the chargino and sneutrino mass. These masses are obtained by diagonalising the corresponding masss matrices which yields the masses as an output. The sneutrino masses, $m_{\tilde{\nu}_i}$, are obtained by diagonalising the sneutrino mass matrix as,

$$\begin{pmatrix} \cos\theta_{\tilde{\nu}} & \sin\theta_{\tilde{\nu}} \\ -\sin\theta_{\tilde{\nu}} & \cos\theta_{\tilde{\nu}} \end{pmatrix} \begin{pmatrix} m^2_{\tilde{L}_{11}} + \mathcal{D}^{\nu}_{L} & m^2_{\tilde{L}_{12}} \\ m^2_{\tilde{L}_{12}} & m^2_{\tilde{L}_{22}} + \mathcal{D}^{\nu}_{L} \end{pmatrix} \begin{pmatrix} \cos\theta_{\tilde{\nu}} & -\sin\theta_{\tilde{\nu}} \\ \sin\theta_{\tilde{\nu}} & \cos\theta_{\tilde{\nu}} \end{pmatrix} = \mathrm{diag}(m^2_{\tilde{\nu}_1}, m^2_{\tilde{\nu}_2}) \quad (6.10)$$

To see how this is done for our work using Python the reader is referred to appendix E. The sneutrino masses are then extracted by square rooting the eigenvalues produced in the calculation as indicated in equation (6.10). Although equation (6.10) indicated the matrix multiplication returns a diagonal result, if one were to perform the matrix multiplication one would find that the result is not in general diagonal. For the above result to satisfy the condition of producing a diagonal result and hence having the sneutrino masses as an output, the mixing angle between the first two generations of sneutrinos is required to have the relation,

$$\tan(2\theta_{\tilde{\nu}}) = \frac{2m^2_{\tilde{L}_{12}}}{m^2_{\tilde{L}_{11}} - m^2_{\tilde{L}_{22}}}. \quad (6.11)$$

Where the above follows the notation that $m_{\tilde{L}_{22}}, m^2_{\tilde{L}_{22}}$ are the flavour diagonal soft masses and the lepton flavour violating mass, $m^2_{\tilde{L}_{12}}$, is defined as a function of the flavour violating parameter $\delta^l_{12}$,

$$m^2_{\tilde{L}_{12}} \equiv (\delta^l_{12})_{LL}\sqrt{m^2_{\tilde{L}_{11}} m^2_{\tilde{L}_{22}}}. \quad (6.12)$$

In equation 6.10, $\mathcal{D}^{\nu}_{L}$ is used to represent the D-terms and for the sneutrinos are defined as [42],

$$\mathcal{D}^{\nu}_{L} = M^2_Z T^v_{L\nu}\cos(2\beta) \quad (6.13)$$

where $M_Z$ is the mass of the Z-boson, $T^v_{L\nu}$ is the third isospin component for the neutrino.

The matrices $U$ and $V$ in equation (6.3) are the eigenvectors of the chargino mass matrix used to produce the required chargino masses. To do this one defines these matrices which diagonalise the chargino mass matrix as,

$$U^* M_{\tilde{\chi}^{\pm}} V^{\dagger} = diag(m_{\tilde{\chi}^{\pm}_1}, m_{\tilde{\chi}^{\pm}_2}). \quad (6.14)$$

The chargino mass matrix is defined in Ref. [42] and produces the chargino

masses required for our calculations when diagonalised,

$$M_{\tilde{\chi}^{\pm}} = \begin{bmatrix} M_2 & \sqrt{2}M_W \sin(\beta) \\ \sqrt{2}M_W \cos(\beta) & \mu \end{bmatrix}. \tag{6.15}$$

where $M_W$ is the mass of the W boson. The angle $\beta$ in equation (6.15) is given by $\tan(\beta) = \frac{v_u}{v_d}$ where $v_{u(d)}$ represents the vacuum expectation value of the Higgs doublet $H_{u(d)}$. Python can calculate the right handed eigenvector of a matrix (see appendix E), so in order to have the eigenvalues (the masses) of the chargino mass matrix isolated on one side, we multiply the equation by its Hermitian conjugate,

$$M_{\tilde{\chi}^{\pm}} X = X diag(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_2^{\pm}}) \tag{6.16}$$

$$\rightarrow X^{\dagger} M_{\tilde{\chi}^{\pm}}^{\dagger} M_{\tilde{\chi}^{\pm}} X = diag(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_2^{\pm}}) X^{\dagger} X diag(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_2^{\pm}}) \tag{6.17}$$

$$= diag(m^2_{\tilde{\chi}_1^{\pm}}, m^2_{\tilde{\chi}_2^{\pm}}) \tag{6.18}$$

one finds [42],

$$V = \begin{bmatrix} \cos(\phi_v) & \sin(\phi_v) \\ -\sin(\phi_v) & \cos(\phi_v) \end{bmatrix}. \tag{6.19}$$

if $\det(M_{\tilde{\chi}^{\pm}}) < 0$ then $V \rightarrow \sigma_3 V$ where $\sigma_3$ is the standard Pauli matrix,

$$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{6.20}$$

with $\phi$ defined as,

$$\tan(\phi_v) = \frac{2\sqrt{2}M_W(\mu \cos\beta + M_2 \sin\beta)}{M_2^2 - \mu^2 + 2M_W^2 \cos 2\beta}. \tag{6.21}$$

However since the matrices that multiply the chargino matrix to diagonalise it are not the same on the left and right handed side it is important to repeat equation (6.17) as if one was multiplying from the left,

$$Y M_{\tilde{\chi}^{\pm}} M_{\tilde{\chi}^{\pm}}^{\dagger} Y^{\dagger} = diag(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_2^{\pm}}) Y Y^{\dagger} diag(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_2^{\pm}}) \tag{6.22}$$

this method is used to obtain the matrix $U$ from equation (6.14),

$$U = \begin{bmatrix} \cos(\phi_u) & \sin(\phi_u) \\ -\sin(\phi_u) & \cos(\phi_u) \end{bmatrix}. \tag{6.23}$$

where,

$$\tan(\phi_u) = \frac{2\sqrt{2}M_W(\mu \sin\beta + M_2 \cos\beta)}{M_2^2 - \mu^2 + 2M_W^2 \cos 2\beta}. \tag{6.24}$$

### 6.1.2 Neutralino and slepton contributions

Calculating the contributions from the neutralinos and charged sleptons is slightly more complicated than the previously discussed chargino case since even only considering the first two generations mixing, the mass matrix is $4 \times 4$ rather than just $2 \times 2$ which makes the calculation for diagonalising the mass matrix to obtain the masses more intensive. The neutralino masses were calculated by diagonalising the following mass matrix,

$$\mathcal{M} = \begin{bmatrix} M_1 & 0 & -M_Z c_\beta s_W & M_Z s_\beta c_W \\ 0 & M_2 & M_Z c_\beta s_W & -M_Z s_\beta c_W \\ -M_Z c_\beta s_W & M_Z s_\beta c_W & 0 & -\mu \\ M_Z c_\beta s_W & -M_Z s_\beta c_W & -\mu & 0 \end{bmatrix}. \tag{6.25}$$

here $M_Z$ is the mass of the Z boson, $c_\beta = \cos(\beta)$ and $c_w = \cos(\theta_w)$. This is diagonalized as,

$$N^* \mathcal{M} N^{-1} = M^{\tilde{x}^0} \tag{6.26}$$

$$M^{\tilde{x}^0 2} = N \mathcal{M}^\dagger \mathcal{M} N^{-1} \tag{6.27}$$

where $N$ is a unitary $4 \times 4$ matrix and $M^{\tilde{x}^0}$ is a diagonal matrix with its elements as the neutralino masses used in the following calculations. Using the above, the neutralino component of $a_\mu$ written as,

$$a_\mu^{\tilde{\chi}_i^0} = \frac{m_\mu}{16\pi^2} \sum_m \Bigg[ -\frac{m_\mu}{12 m^2_{\tilde{\chi}_i^0}} [n^{L*}_{\mu i m} n^L_{\mu i m} + n^R_{\mu i m} n^{R*}_{\mu i m}] x_{im} F_1^N(x_{im}) + \frac{1}{3 m_{\tilde{\chi}_i^0}} Re[n^L_{\mu i m} n^R_{\mu i m}] x_{im} F_2^N(x_{im}) \Bigg] \tag{6.28}$$

here $F_1^N$, $F_2^N$ and $n^L_{\ell i m}$ are loop functions defined in [23] with $x_{im} \equiv \frac{m^2_{\tilde{\chi}_i^0}}{m^2_{\tilde{\ell}_m}}$,

$$n^L_{\ell i m} = \frac{1}{\sqrt{2}}(g_1 N_{i1} + g_2 N_{i2}) K^*_{m,\ell} - Y_\ell N_{i3} K^*_{m,\ell+3}, \tag{6.29}$$

$$n^R_{\ell i m} = \sqrt{2} g_1 N_{i1} K_{m,\ell+3} + Y_\ell N_{i3} K_{m,l}, \tag{6.30}$$

$$F_1^N(x) = \frac{2}{(1-x)^4}\left[1 - 6x + 3x^2 + 2x^3 - 6x^2 \ln x\right], \tag{6.31}$$

$$F_2^N(x) = \frac{3}{(1-x)^3}\left[1 - x^2 + 2x \ln x\right], \tag{6.32}$$

$Y_\ell$ and $g_i$ are the Yukawa and gauge couplings respectively and follow equations (6.5) and (6.6). The matrices $K$ and $N$ are defined as the matrices which diagonalise the slepton and neutralino mass matrices respectively to produce the corresponding masses as eigenvalues. The slepton masses are calculated by,

$$K\mathcal{M}^2 K^\dagger = diag(m^2_{\tilde{\ell}_1}, ..., m^2_{\tilde{\ell}_6}), \tag{6.33}$$

where $\mathcal{M}$ is the slepton mass matrix which has the form,

$$\mathcal{M}^2 = \begin{bmatrix} m^2_{\tilde{L}_{11}} + \mathcal{D}_{L,e} & 0 & 0 & m^{2\dagger}_{LR_e} & 0 & 0 \\ 0 & m^2_{\tilde{L}_{22}} + \mathcal{D}_{L,\mu} & 0 & 0 & m^{2\dagger}_{LR_\mu} & 0 \\ 0 & 0 & m^2_{\tilde{L}_{33}} + \mathcal{D}_{L,\tau} & 0 & 0 & m^{2\dagger}_{LR_\tau} \\ m^2_{LR_e} & 0 & 0 & m^2_{\tilde{R}_{11}} + \mathcal{D}_{R,e} & 0 & 0 \\ 0 & m^2_{LR_\mu} & 0 & 0 & m^2_{\tilde{R}_{22}} + \mathcal{D}_{R,\mu} & 0 \\ 0 & 0 & m^2_{LR_\tau} & 0 & 0 & m^2_{\tilde{R}_{33}} + \mathcal{D}_{R,\tau} \end{bmatrix}. \tag{6.34}$$

where $m_{\tilde{L}_{11}}, m_{\tilde{L}_{22}}, m_{\tilde{L}_{33}}, m_{\tilde{R}_{11}}, m_{\tilde{R}_{22}}$ and $m_{\tilde{R}_{33}}$ are the independent soft masses. The off-diagonal elements are given by $(m^2_{LR})_{ij} = \delta_{ij}(A_i - \mu^* \tan\beta)$ where $A_i$ are the soft mass trilinear couplings. The eigenvectors $K$ and $N$ used in our calculations are calculated numerically in Python using the method outlined in appendix E and as such the analytic expressions for these are not used and hence omitted.

## 6.2 $B(\mu \to e\gamma)$ in the Minimally Supersymmetric Standard Model

The source of LFV in the MSSM is the misalignment of the slepton and lepton mass matrices arising from the off diagonal mass parameters seen in the slepton mass matrix, $m_{\tilde{L}_{ij}}$ where $i \neq j$ in equation (6.10). The LFV $\mu - e$ process in the MSSM arises from the mass $m_{\tilde{L}_{12}}$ and the branching ratio for

the decay $\mu \to e\gamma$ is calculated by examining Feynman diagrams that are very similar to those for $a_\mu$. The resulting expression for this branching ratio is,

$$\mathrm{BR}(\mu \to e\gamma) = \frac{3\pi^2 e^2}{G_F^2 m_\mu^4} \left(|a_{\mu e\gamma L}|^2 + |a_{\mu e\gamma R}|^2\right). \tag{6.35}$$

The similarity of the Feynman diagrams leads to the same ability to decompose the amplitudes into chargino and sneutralino components,

$$a_{mue\gamma L} = \sum_k a^{\tilde{\chi}_k^\pm}_{\mu e\gamma L} + \sum_i a^{\tilde{\chi}_i^0}_{\mu e\gamma L}. \tag{6.36}$$

However for the right handed part the amplitude is just given by the sneutralino as demonstrated in [86],

$$\sum_k a^{\tilde{\chi}_k^\pm}_{\mu e\gamma R} = 0 \tag{6.37}$$

### 6.2.1 Chargino Component to $B(\mu \to e\gamma)$

The chargino contribution to the left handed amplitude is given by,

$$a^{\tilde{\chi}_k^\pm}_{\mu e\gamma L} = a^{\tilde{\chi}_k^\pm}_{\mu e\gamma\ 1} + a^{\tilde{\chi}_k^\pm}_{\mu e\gamma\ 2}, \tag{6.38}$$

where $I$ is the chargino index. From Ref. [23] ratio of the left handed amplitude and the contribution to $a_\mu$ in the chargino sector can be written as,

$$\frac{a^{\tilde{\chi}_k^\pm}_{\mu e\gamma\ I}}{a^{\tilde{\chi}_k^\pm}_{\mu\ I}} = \frac{\sin 2\theta_{\tilde{\nu}}}{2} \frac{x_{k1} F_I^C(x_{k1}) - x_{k2} F_I^C(x_{k2})}{\sin^2\theta_{\tilde{\nu}}\, x_{k1} F_I^C(x_{k1}) + \cos^2\theta_{\tilde{\nu}}\, x_{k2} F_I^C(x_{k2})}\ , \quad I = 1, 2. \tag{6.39}$$

which allows the calculation of $a^{\tilde{\chi}_k^\pm}_{\mu e\gamma L}$ since $a^{\tilde{\chi}_k^\pm}_{\mu}$ is calculated as discussed previously in section 6.1.1.

### 6.2.2 Neutralino Component to $B(\mu \to e\gamma)$

For the neutralino contribution to the amplitude neither the left nor right handed components are zero in general.

$$\frac{16\pi^2}{m_\mu} a^{\tilde{\chi}_i^0}_{\mu e\gamma R} = \sum_m \left[ -\frac{m_\mu}{12 m^2_{\tilde{\chi}_i^0}} n^R_{\mu im} n^{R*}_{eim} x_{im} F_1^N(x_{im}) + \frac{1}{3 m_{\tilde{\chi}_i^0}} n^{L*}_{\mu im} n^{R*}_{eim} x_{im} F_2^N(x_{im}) \right],$$
$$\frac{16\pi^2}{m_\mu} a^{\tilde{\chi}_i^0}_{\mu e\gamma L} = \sum_m \left[ -\frac{m_\mu}{12 m^2_{\tilde{\chi}_i^0}} n^{L*}_{\mu im} n^{L}_{eim} x_{im} F_1^N(x_{im}) + \frac{1}{3 m_{\tilde{\chi}_i^0}} n^{R}_{\mu im} n^{L}_{eim} x_{im} F_2^N(x_{im}) \right]. \tag{6.40}$$

where the loop functions are the same as for the neutralino component to $a_\mu$ given by equations (6.29) to (6.32).

These two components then allow one to calculate the branching ratio for $\mu \to e\gamma$ given in equation (6.35). One can see from the relevant equations the large overlap in the parameters used in the calculation of this LFV observable and $a_\mu$, an interesting relation mentioned at the start of the chapter.

## 6.3 Additional Lepton Flavour Violation in the Minimally Supersymmetric Standard Model

As discussed earlier there are motivations for studying $\mu - \tau$ LFV in the MSSM and the impact these have on the parameter space. To consider LFV involving $\tau$ decays one must consider the third generation soft breaking masses,

$$m_{\tilde{L}_{33}} \neq 0 \tag{6.41}$$

$$m_{\tilde{R}_{33}} \neq 0 \tag{6.42}$$

as well as the off diagonal flavour violating mass,

$$m_{\tilde{L}_{23}} = \sin(2\theta_L)\sqrt{m^2_{\tilde{L}_{22}} - m^2_{\tilde{L}_{33}}}, \tag{6.43}$$

with $\theta_L$ denoting the mixing angle. Similarly the mass $m_{\tilde{R}_{23}}$ is defined as in equation (6.43) but with $L \to R$. This off diagonal mass is responsible for inducing the following LFV observables relating to $\tau$ decays [47],

$$BR(\tau^- \to \mu^- \gamma) = \frac{48\pi^3 \alpha_{em}}{G_F^2} \left[ |D_L^\gamma|^2 + |D_R^\gamma|^2 \right] BR(\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau) \tag{6.44}$$

where $\alpha_{em} = e^2/(4\pi) \simeq 1/137$, $G_F$ is the Fermi constant and $BR(\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau)$ is taken to be 17.39% [17] and the additional terms are defined in appendix G along with those terms contributing to the other LFV decays mentioned below. The introduction of these discussed third generation masses also gives rise to $\tau$ decays of the form $\tau \to lll$ which constrain the parameter space further [47],

$$\begin{aligned} BR(\tau^- \to \mu^- e^+ e^-) = \frac{1}{8G_F^2} \Big[ & |F_L^{e_L}|^2 + |F_L^{e_R}|^2 + |F_R^{e_L}|^2 + |F_R^{e_R}|^2 \\ & + 4e^2 \mathrm{Re}\left( D_L^\gamma (\bar{F}_L^{e_L} + \bar{F}_L^{e_R}) + D_R^\gamma (\bar{F}_R^{e_L} + \bar{F}_R^{e_R}) \right) \\ & + 8e^4 (|D_L^\gamma|^2 + |D_R^\gamma|^2) \left( \log \frac{m_\tau^2}{m_e^2} - 3 \right) \Big] BR(\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau), \end{aligned} \tag{6.45}$$

$$\begin{aligned} BR(\tau^- \to \mu^- \mu^+ \mu^-) = \frac{1}{8G_F^2} \Big[ & 2|F_L^{\mu_L}|^2 + |F_L^{\mu_R}|^2 + |F_R^{\mu_L}|^2 + 2|F_R^{\mu_R}|^2 \\ & + 4e^2 \mathrm{Re}\left( D_L^\gamma (2\bar{F}_L^{\mu_L} + \bar{F}_L^{\mu_R}) + D_R^\gamma (\bar{F}_R^{\mu_L} + 2\bar{F}_R^{\mu_R}) \right) \\ & + 8e^4 (|D_L^\gamma|^2 + |D_R^\gamma|^2) \left( \log \frac{m_\tau^2}{m_\mu^2} - \frac{11}{4} \right) \Big] BR(\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau). \end{aligned} \tag{6.46}$$

These decays are forbidden in the Standard Model, however the non-zero experimental bounds on these observables offer a possible indicator to physics beyond the Standard Model. The upper limits on these $\tau-$decays are,

$$Br(\tau \to \mu\gamma) < 1.3 \times 10^{-9} \text{ [48]}, \tag{6.47}$$

$$Br(\tau \to \mu ee) < 1.8 \times 10^{-8} \text{ [48]}, \tag{6.48}$$

$$Br(\tau \to 3\mu) < 2.1 \times 10^{-8} \text{ [49]}, \tag{6.49}$$

which constrain the parameter space by discarding data sets which produce values above these limits.

## 6.4 Minimally Supersymmetric Standard Model Input

It was first decided to check the framework from this project was capable of reproducing results from literature to ensure that any further analysis was build upon a solid foundation. The following plot of $Br(\mu \to e\gamma)$ vs $a_\mu$ was taken from [23] since the assumptions involved were minimal and could easily be lifted for further analysis,

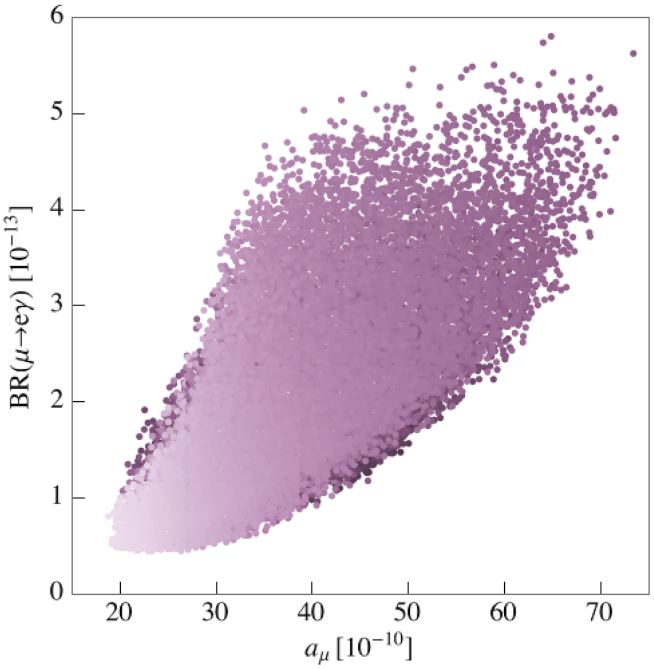



**Figure 6.1:** *$Br(\mu \to e\gamma)$ vs $a_\mu$ from [23].*

which involved sampling the masses $M_1$, $M_2$, $\mu$, $m_{\tilde{L}_{11}}$, $m_{\tilde{L}_{22}}$, $m_{\tilde{R}_{11}}$, $m_{\tilde{R}_{22}}$, independently between 300 and 600 GeV whilst fixing $\tan(\beta) = 50$ and the lepton flavour violating parameters $(\delta^l_{12})_{LL} = (\delta^l_{12})_{RR} = 2 \times 10^{-5}$ as specified in [23]. The scatter plots shown in figures 6.1 and 6.2 show the correlation between $a_\mu$ and $Br(\mu \to e\gamma)$ with the parameter space corresponding to the maximum value of branching ratio for $\mu \to e\gamma$ at the time of $5.7 \times 10^{-13}$ correlating to higher values of $a_\mu$ which are indeed excluded by experiment. An important note is these scatter plots do not show the likelihood of each value and so the fact that some values of $a_\mu$ plotted are excluded by experiment is not reflected in these plots, only impact of both of these observables on constraining the MSSM since higher values of $a_\mu$ correspond to higher values of

$Br(\mu \to e\gamma)$.

Although more recent measurements of $a_\mu$ have been undertaken, for consistency we used the measured value at the time of publication of [23] of $a_\mu = (287 \pm 80) \times 10^{-11}$. This value is then used to construct a Gaussian likelihood in the form of equation (2.7) with $\mu = 287 \times 10^{-11}$ and $\sigma = 80 \times 10^{-11}$,

$$P(x)_{a_\mu} = \frac{1}{\sqrt{2\pi(80 \times 10^{-11})^2}} e^{-\frac{(x-287\times10^{-11})^2}{2(80\times10^{-11})^2}}, \tag{6.50}$$

where $x$ denotes the value of $a_\mu$ calculated in our framework using the parameters sampled to calculate the relevant components from sections 6.1.1 and 6.1.2.

For other observables which can be fitted to an n-dimensional Gaussian distribution (equation 2.9) which was done for the results in chapter 5, however not all observables can be fitted this way. For observables such as the branching ratios for lepton flavour violating decays only an upper bound is known, so for this we treat it as a flat distribution,

$$P(x)_{\text{flat}} = \begin{cases} \frac{1}{b-a} & \forall x \in [a, b], \\ 0 & \text{otherwise} \end{cases} \tag{6.51}$$

where $b$ and $a$ are the boundaries of the distribution such that $b < a$. For the branching ratio we have the condition $Br(\mu \to e\gamma) < 5.7 \times 10^{-13}$ at the time of publishing of [23]. In this case $b = 5.7 \times 10^{-13}$ and $a = 0$ so we have a likelihood for it,

$$P(x)_{Br(\mu\to e\gamma)} = \begin{cases} \frac{1}{5.7\times10^{-13}} & \forall x \in [0, 5.7 \times 10^{-13}], \\ 0 & \text{otherwise.} \end{cases} \tag{6.52}$$

It is important to note that we are not concerned with any normalizations here as our analysis focuses on calculating the Bayes factor between two models with the same observables so any normalization constants cancel out in the ratio. As shown in section 6.2, the branching ratio for $\mu \to e\gamma$ is broken down into the chargino and neutralino components (sections 6.2.1 and 6.2.2 respectively) and summed over for the total. One can see from equation 6.52 that any value we calculate for $Br(\mu \to e\gamma)$ will have the same likelihood value which reflects the fact that we do not have any information to indicate which values of $Br(\mu \to e\gamma)$ within the range would be favoured. At first glance one may be

tempted to conclude that this adds no statistical information to the model and as such the Bayes factor between two models containing only such observables would be equal to one, however this is not the case. It is important to remember that observables such as $Br(\mu \to e\gamma)$ only have a non-zero likelihood when the calculated value is below the experimentally identified upper bound, so a likelihood of the form given in equation 6.52 represents the likelihood that a sampled parameter set produces a value for $Br(\mu \to e\gamma) < 5.7 \times 10^{-13}$. So if one were to calculate the Bayes factor of a model containing only observables of this kind, the divergence from a Bayes factor equal to one would mean that one models parameter space would be more likely to produce an allowed value of the observable.

Since $a_\mu$ and $Br(\mu \to e\gamma)$ are uncorrelated the likelihood for the model can easily be calculated by simply multiplying the individual likelihoods together,

$$P(x)_{\text{model}} = P(x)_{a_\mu} P(x)_{Br(\mu \to e\gamma)}, \tag{6.53}$$

so for our test we employ Markov-chain Monte-Carlo sampling for our parameters, the method of which is outlined in section 4.2 where our $g(x)$ here (see equation 4.2) is the likelihood specified above in equation 6.53. This then allows us to reproduce figure 6.1,

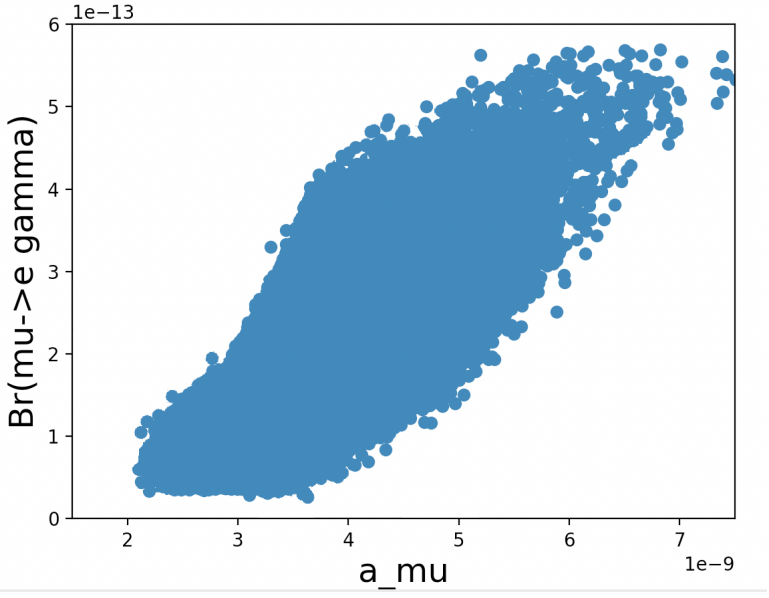



**Figure 6.2:** *$Br(\mu \to e\gamma)$ vs $a_\mu$ following the same constraints as figure 6.1 outlined previously.*

Figures 6.1 and 6.2 above demonstrate consistency with the literature and results produced in our computer framework subject to the same constraints. To calculate the Bayes factor between two models it is important to be as general as possible with the model parameters to prevent the result being influenced by the choice of a fixed parameter value. The previous result was concerned with only the first two generations of sleptons which involved ignoring the third generation soft masses $m_{\tilde{L}_{33}}$ and $m_{\tilde{R}_{33}}$, fixing $\tan(\beta) = 50$ and setting the lepton flavour violating parameters, $(\delta^l_{12})_{LL}, (\delta^l_{12})_{RR}$ as $2 \times 10^{-5}$. Since this project is concerned with comparing models in as general form as possible it made sense to relax these fixtures and consider their impact. Experimental and theoretical bounds on parameters are useful for determining the ranges to sample over, however this kind of information is not always available. For cases where there is no data to inform the bounds, the sample range is taken to be somewhat arbitrary whilst still making sense from a physics point of view. For example, in the case of the branching ratio $Br(\mu \to e\gamma)$ which has an experimental bound of $Br(\mu \to e\gamma) < 4.2 \times 10^{-13}$ the parameters have to reflect this and their ranges should be chosen to keep in line with this bound on the observable. Of course this does not have to be exact because the Metropolis-Hastings once it reaches equilibrium will sample the values according to the target distribution, equation (6.52) for $Br(\mu \to e\gamma)$, which means the sampling will occur in regions where the calculated observable value obey this distribution. However selecting a parameter range too broad could add to the required burn-in length (see section 4.2) especially if one is considering multiple parameters with extremely broad sample ranges. These are considerations taking into account for the next part of the analysis.

To consider as broad a range as possible for the MSSM we remove the restrictions discussed and sample the masses in the following range,

$$M_{\text{SUSY}} \leq 1 \text{ TeV} \quad |\mu| \leq 3 \text{ TeV}$$

these ranges were selected as values greater than this range lead to calculations of $\Delta a_\mu^{2L}$ larger than the measured difference [45] reported as,

$$a_\mu^{EXP} - a_\mu^{SM} = (268 \pm 76) \times 10^{-11}$$

The value for $\tan(\beta)$ is restricted because at large values ($> 50$) the fermion

masses are affected by (forbidden) couplings to the so called "wrong" Higgs doublet [44]. One can see this with the bottom quark mass,

$$m_b = y_b v_d + \delta\tilde{y}_b v_u, \tag{6.54}$$

where $v_u$ and $v_d$ represent the VeVs of the Higgs doublets $H_u$ and $H_d$ respectively with $\tan(\beta) = \frac{v_u}{v_d}$ and $y_b$ is the Yukawa coupling and $\delta\tilde{y}_b$ is the Yukawa coupling to the "wrong" (up-type) Higgs doublet. To see the dependence of $m_b$ on $\tan(\beta)$ one can rewrite equation (6.54) as,

$$m_b = y_b v_d + (1 + \frac{\delta\tilde{y}_b}{y_b}\tan(\beta)). \tag{6.55}$$

This allows one to see that for large $\tan(\beta)$ the mass of the bottom quark can exceed that of the top given by, $m_t = y_t v_u$ and so to preserve the mass hierarchy of the Standard Model masses, we avoid large $tan(\beta)$ which gives a range of values,

$$1 \leq \tan(\beta) \leq 50 \tag{6.56}$$

Finally the lepton flavour violating parameter $(\delta^l_{12})_{LL}$ is varied broadly enough as to not introduce a bias in results whilst making sure not to violate the experimental constraints on $Br(\mu \to e\gamma)$ [45]. This gives a range for us to sample over of,

$$-7.5 \times 10^{-5} < (\delta^l_{12})_{LL}, (\delta^l_{12})_{RR} < 7.5 \times 10^{-5}. \tag{6.57}$$

To calculate the LFV observables relating to the $\tau-$decay in the MSSM one must not only consider the flavour violating masses $m_{\tilde{L}_{12}}$ and $m_{\tilde{R}_{12}}$ but also $m_{\tilde{L}_{23}}$ and $m_{\tilde{R}_{23}}$. Equation (6.43) gives these masses as functions of the mixing angle $\theta_L$ and $\theta_R$ respectively. Mathematically of course the term $\sin(2\theta_{\{L,R\}})$ has a range $[-1, 1]$ however for a positive flavour changing mass the region is restricted to only positive values which give the following bound on the angles,

$$0 \leq \theta_{\{L,R\}} \leq \frac{\pi}{4}. \tag{6.58}$$

The trilinear couplings $A_i$ mentioned in equation (6.34) have been shown to not impact on the calculation of $a_\mu$ and $Br(\mu \to e\gamma)$ [23]. However when calculating LFV decays involving the $\tau$ then the couplings $A_\tau, A_\mu$ and $A^{\{L,R\}}_{\mu\tau}$ can have significant impact on the results. The trilinear couplings have a limit

of,

$$|A_{\{\tau,\mu,\tau\mu,\mu\tau\}}| \leq M_{\text{SUSY}}, \tag{6.59}$$

as values outside of this range can lead to destabilizing of the scalar potential as well as the introduction of VeVs for the slepton fields [47].

All of the discussion above regarding the parameters and their ranges is summarized in table 6.1. With the prior probability representing the distribution they were sampled according to and since these are sampled uniformly over their ranges since there is no direct experimental measurements, meaning no mean values or standard deviations, the prior is taken to be a uniform distribution as to not introduce bias in the results,

$$P(x)_{\text{flat}} = \begin{cases} \frac{1}{b-a} \ \forall x \in [a,b], \\ 0 \quad \text{otherwise} \end{cases} \tag{6.60}$$

To study the parameter space of the model one must construct the likelihood that the sample points must be fitted to in the Markov-chain, $g()$ in equation (4.2). The observables discussed in this model were $a_\mu$, $Br(\mu \to e\gamma)$, $Br(\tau \to \mu\gamma)$, $Br(\tau \to 3\mu)$ and $Br(\tau \to \mu ee)$. As discussed in the initial part of our analysis for the MSSM, the LFV branching ratios are only measured to an upper bound, there is currently no measured mean value so as with $Br(\mu \to e\gamma)$ they are assigned a uniform distribution as given in equation (6.51) and since the likelihoods of the observables are uncorrelated then the total likelihood for the model is just the product of the individual distributions (as with equation (6.53)). This allows the construction of the likelihood for the model which serves as the target distribution for the sampling,

$$\begin{aligned} P(D|\vec{\theta}_{\text{MSSM}}) = P(a_\mu|\vec{\theta}_{\text{MSSM}})P(Br(\mu \to e\gamma)|\vec{\theta}_{\text{MSSM}})P(Br(\tau \to \mu\gamma)|\vec{\theta}_{\text{MSSM}}) \\ \times P(Br(\tau \to 3\mu)|\vec{\theta}_{\text{MSSM}})P(Br(\tau \to \mu ee)|\vec{\theta}_{\text{MSSM}}). \end{aligned} \tag{6.61}$$

with the likelihood for $a_\mu$ given by equation (6.50) however as discussed the value of the experimental value is update to $(268 \pm 76) \times 10^{-11}$ to reflect more recent work done to refine the value since the publication of Ref. [23]. The

| Parameter | Range | Prior, $P(\theta_i\|M)$ |
|---|---|---|
| $M_1$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $M_2$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{L}_{11}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{L}_{22}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{R}_{11}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{R}_{22}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{L}_{33}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $m_{\tilde{R}_{33}}$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\mu$ | $[-3000, 3000]$ GeV | $\begin{cases} \frac{1}{6000} \ \forall\theta_i \in [0, 6000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\tan(\beta)$ | $[1, 50]$ | $\begin{cases} \frac{1}{49} \ \forall\theta_i \in [0, 50], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $(\delta^l_{12})_{LL}$ | $[0, 7.5 \times 10^{-5}]$ | $\begin{cases} \frac{1}{7.5\times10^{-5}} \ \forall\theta_i \in [0, 7.5 \times 10^{-5}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\theta_L$ | $[0, \frac{\pi}{4}]$ | $\begin{cases} \frac{4}{\pi} \ \forall\theta_i \in [0, \frac{\pi}{4}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\theta_R$ | $[0, \frac{\pi}{4}]$ | $\begin{cases} \frac{4}{\pi} \ \forall\theta_i \in [0, \frac{\pi}{4}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $A_\tau$ | $[-1000, 1000]$ GeV | $\begin{cases} \frac{1}{2000} \ \forall\theta_i \in [-1000, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $A_\mu$ | $[-1000, 1000]$ GeV | $\begin{cases} \frac{1}{2000} \ \forall\theta_i \in [-1000, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $A_{\mu\tau}$ | $[-1000, 1000]$ GeV | $\begin{cases} \frac{1}{2000} \ \forall\theta_i \in [-1000, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $A_{\tau\mu}$ | $[-1000, 1000]$ GeV | $\begin{cases} \frac{1}{2000} \ \forall\theta_i \in [-1000, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |

**Table 6.1:** *Table showing the parameters used in our calculations in the MSSM along with the ranges they were sampled over and their prior distribution where $\theta_i$ denotes the sampled value for the parameter.*

LFV decays having likelihoods of the form,

$$P(Br(\mu \to e\gamma)|\vec{\theta}) = \begin{cases} \frac{1}{4.2\times10^{-13}} \ \forall\kappa \in [0, 4.2 \times 10^{-13}], \\ 0 \qquad\qquad \text{otherwise.} \end{cases} \tag{6.62}$$

here $\kappa$ denotes the value of the observable $Br(\mu \to e\gamma)$ calculated with the set of sampled parameters $\vec{\theta}$. For the other LFV observables $4.2 \times 10^{-13}$ is replaced with the relevant experimental upper bound for the decay. This likelihood is then used as a target distribution in which one can then study the parameter space allowed for all such observables to occur.

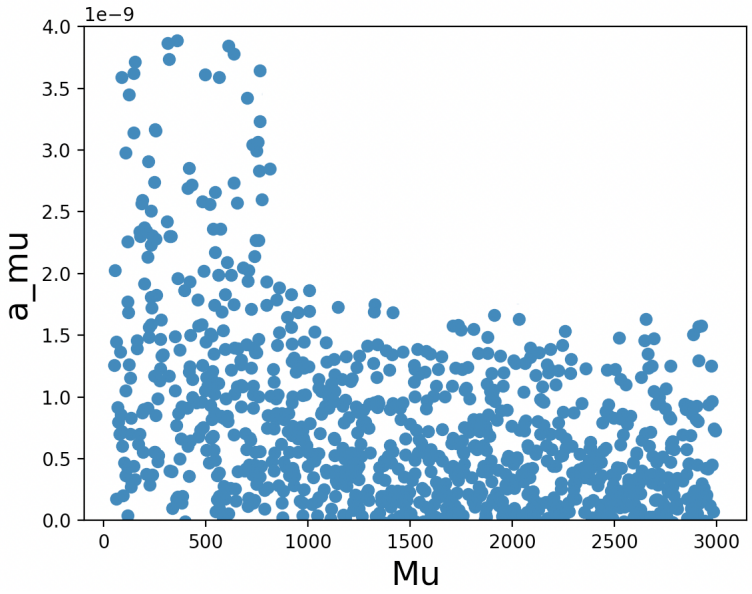



**Figure 6.3:** *Plot showing* $10^5$ *points of* $\mu$ *in GeV against* $a_\mu$ .

Figure 6.3 shows favoured values of $\mu$ as being positive. Values of $\mu > 0$ against $a_\mu$ are expected to be preferred since negative values for the Higgsino mass parameter result in $a_\mu < 0$ which is strongly disfavoured by experiment.

Due to the large number of parameters we freely sample over this often leads to no obvious preferred values of each parameter as the effect of one of these parameters on the calculation can easily be offset by one of the other multiplitude of parameters in the model, the exception being $\mu$ whose value being negative leads explicitly to values of $a_m u < 0$ which are strongly disfavoured by experiment.

This chapter has set out the fundamentals of the MSSM and the parameters required for the observables to be calculated which then allows the marginalized likelihood to be calculated for the Bayes factor.

# Chapter 7

# Two-Higgs Doublet Model

The 2HDM is an extension of the Standard Model achieved by the addition of an extra scalar doublet with the same hypercharge as the first. With the discovery of the Higgs boson it is important to consider whether this is a unique boson as predicted by the Standard Model or whether there exist other bosonic particles contributing to electroweak symmetry breaking.

The basics of the 2HDM are set out initially in this chapter along with the parameters used in the calculations in this model. The anomalous magnetic moment of the muon, $a_\mu$ in the 2HDM followed by its decomposition into the one loop contributions plus the two loop Fermionic and bosonic components with the method for calculating these elements presented with the relevant loop integrals. As discussed previously with the MSSM in chapter 6 the similarity in the Feynman diagrams makes studying the LFV decay $\mu \to e\gamma$ along with $a_\mu$ an interesting exercise with the decay introducing additional parameters in the form of the flavour violating Yukawa couplings $\rho_e^{e\mu,\mu e}$ with usual analysis of this decay involving setting all other Yukawa couplings to zero. As shown later in the chapter, the calculation for $Br(\mu \to e\gamma)$ in the 2HDM involves the sum of the one-loop contribution with the two-loop contribution from Barr-Zee diagrams (see Fig. 7.1) and is identical to the calculation for $Br(\tau \to \mu\gamma)$ except for the lepton masses being a generation higher, $m_e \to m_\mu$ and $m_\mu \to m_\tau$.

When calculating the Bayes factor between the two models it is important to remember that for a meaningful comparison, both models must fit the same data meaning that the observables considered in the 2HDM are the

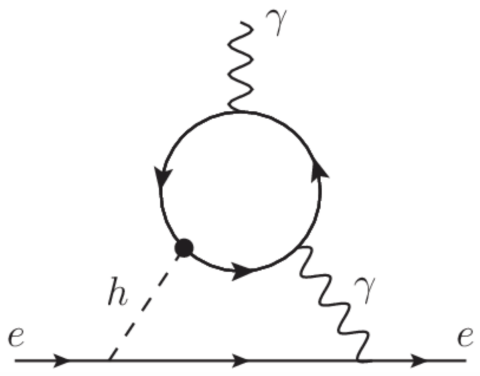



**Figure 7.1:** *Example of a Barr-Zee diagram [50]. The type of which are used in the calculations for the LFV decays $\mu \to e\gamma$ and $\tau \to \mu\gamma$.*

same as those in the MSSM. LFV $\mu - \tau$ decays in the 2HDM can enhance the calculation of $a_\mu$ offering additional links between $a_\mu$ and LFV observables making them an important topic to study together in the 2HDM. As with the MSSM, the effects of these LFV observables and $a_\mu$ on the parameter space is examined with the aim of constructing a 2HDM that can accommodate these deficiencies of the Standard Model. The constrained parameter space of this 2HDM is shown and the regions of allowed parameters discussed.

As mentioned the 2HDM consists of two Higgs doublets,

$$\phi_i = \begin{pmatrix} a_i^+ \\ \frac{1}{\sqrt{2}}(v_i + b_i + ic_i) \end{pmatrix}, \quad i = 1, 2. \tag{7.1}$$

The potential for this model is [51],

$$\begin{aligned} V(\phi_1, \phi_2) = {} & m_{11}^2 \phi_1^\dagger \phi_1 + m_{22}^2 \phi_2^\dagger \phi_2 - \{m_{12}^2 (\phi_1^\dagger \phi_2 + \phi_2^\dagger \phi_1)\} \\ & + \frac{\lambda_1}{2}(\phi_1^\dagger \phi_1)^2 + \frac{\lambda_2}{2}(\phi_2^\dagger \phi_2)^2 + \lambda_3 (\phi_1^\dagger \phi_1 \phi_2^\dagger \phi_2) \\ & + \lambda_4 (\phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1) + \frac{1}{2}(\lambda_5 (\phi_1^\dagger \phi_2)^2 + \\ & [\lambda_6 (\phi_1^\dagger \phi_1) + \lambda_7 (\phi_2^\dagger \phi_2)] \phi_1^\dagger \phi_2 + h.c.) \end{aligned} \tag{7.2}$$

Many times when analysing the two-Higgs doublet model $\lambda_6$ and $\lambda_7$ are zero by application of the symmetry $\Phi_1 \to -\Phi_1$. This symmetry however results in a zero $m_{12}^2$ unless one allows for soft-breaking of $Z_2$ symmetry [51].

$$\begin{pmatrix} \Phi_v \\ \Phi_\perp \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{7.3}$$

Here $\tan(\beta) = \frac{v_2}{v_1}$ is the ratio of vacuum expectation values. The standard model VEV is recovered using $v^2 = v_1^2 + v_2^2$. $\Phi_v$ and $\Phi_\perp$ are given by,

$$\Phi_v = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + S_1 + i\,G^0) \end{pmatrix}, \qquad \Phi_\perp = \begin{pmatrix} H^+ \\ \frac{1}{\sqrt{2}}(S_2 + i\,A) \end{pmatrix}. \tag{7.4}$$

where $H^\pm$ and $A$ correspond to the charged Higgs bosons and the neutral CP-odd psuedoscalar Higgs respectively. $S_1$ and $S_2$ do not directly denote mass states however by introduction of the mixing angle $\alpha$ they can be rotated to give the CP-even neutral Higgs bosons, $H$ and $h$,

$$\begin{pmatrix} H \\ h \end{pmatrix} = \begin{pmatrix} \cos(\beta-\alpha) & -\sin(\beta-\alpha) \\ \sin(\beta-\alpha) & \cos(\beta-\alpha) \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}. \tag{7.5}$$

In the Standard Model $\beta - \alpha = \frac{\pi}{2}$ however experimental results from the LHC allow a deviation defined by,

$$\eta = \frac{\pi}{2} - (\beta - \alpha). \tag{7.6}$$

Equation (7.2) introduces the parameters, $m_{11}^2, m_{22}^2, m_{12}^2, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$. [85] which we replace with the following physical parameters, the Higgs masses $M_h, M_H, M_A, M_{H^\pm}$, the mixing angles $\alpha, \beta$ and the vacuum expectation value $v$, using [51],

$$\lambda_1 = \frac{M_H^2\cos^2(\alpha) + M_h^2\sin^2(\alpha) - m_{12}^2\tan(\beta)}{v^2\cos^2(\beta)}, \tag{7.7}$$

$$\lambda_2 = \frac{M_H^2\sin^2(\alpha) + M_h^2\cos^2(\alpha) - m_{12}^2\cot(\beta)}{v^2\sin^2(\beta)}, \tag{7.8}$$

$$\lambda_3 = \frac{(M_H^2 - M_h^2)\cos(\alpha)\sin(\alpha) + 2M_{H^\pm}^2\sin(\beta)\cos(\beta) - m_{12}^2}{v^2\sin(\beta)cos(\beta)}, \tag{7.9}$$

$$\lambda_4 = \frac{(M_A^2 - 2M_{H^\pm}^2)\sin(\beta)\cos(\beta) + m_{12}^2}{v^2\sin(\beta)\cos(\beta)}, \tag{7.10}$$

$$\lambda_5 = \frac{m_{12}^2 - M_A^2\sin(\beta)\cos(\beta)}{v^2\sin(\beta)\cos(\beta)}. \tag{7.11}$$

The Yukawa couplings, $Y_f^S$ are defined as follows [51],

$$\begin{aligned} Y_f^h &= \sin(\beta-\alpha) + \cos(\beta-\alpha)\zeta_f, \\ Y_f^H &= \cos(\beta-\alpha) - \sin(\beta-\alpha)\zeta_f, \\ Y_{d,l}^A &= -\zeta_{d,l}, \quad Y_u^A = \zeta_u, \end{aligned} \tag{7.12}$$

with the form of $\zeta_f$ depending on which type of 2HDM we are looking at. In our analysis we consider a type-III 2HDM in which both Higgs doublets couple to all the fermions. Below is a table showing the form of $\zeta_f$ depending on which type of 2HDM is being considered. For the type-III scenario $\zeta_f$ is the same as for type-II [52].

| | Type I | Type II | Type X | Type Y |
|---|---|---|---|---|
| $\zeta_u$ | $\cot\beta$ | $\cot\beta$ | $\cot\beta$ | $\cot\beta$ |
| $\zeta_d$ | $\cot\beta$ | $-\tan\beta$ | $\cot\beta$ | $-\tan\beta$ |
| $\zeta_l$ | $\cot\beta$ | $-\tan\beta$ | $-\tan\beta$ | $\cot\beta$ |

**Table 7.1:** *Table showing the relation between the parameters $\zeta_u, \zeta_d$ and $\zeta_l$ to the angle $\beta$ depending on the type of 2HDM being considered.*

## 7.1 $a_\mu$ in the 2HDM

The anomalous magnetic of the muon is a possible indicator of physics beyond the standard model as discussed before. One possible extension that offers contributions toward explaining this anomaly is the two-Higgs doublet model. This deviation has been explored in the lepton specific type-X 2HDM in Refs[78, 79, 80, 81, 82, 83] however this framework was built to accommodate as general a model as possible and as such the calculations relating to $a_\mu$ were input in a such a way that the type of 2HDM one considers only changes the form of $\zeta_f$. The two-loop contributions to $a_\mu$ are more dominant since the one loop result is suppressed by powers of the muon Yukawa coupling. The expressions for these calculations were given in Ref. [51] in full where previously the two-loop Barr-Zee diagrams [53] were calculated in Ref. [54].

The full two-loop result for the 2HDM contribution to $a_\mu$ is given by the sum of the one-loop contributions with the two-loop bosonic and fermionic

contributions. The 2HDM contribution to $a_\mu$ is broken down as,

$$a_\mu^{2HDM} = a_\mu^{2HDM,1-loop} + a_\mu^{\mathrm{B}} + a_\mu^{\mathrm{F}} + a_\mu^{\Delta r\text{-shift}}, \tag{7.13}$$

where $a_\mu^{2HDM,1-loop}$ is the one loop contribution, $a_\mu^{\mathrm{B}}$ is the bosonic contribution, $a_\mu^{\mathrm{F}}$ is the fermionic component and $a_\mu^{\Delta r\text{-shift}}$ is the component relating to $\Delta r$ which relates the electroweak gauge boson masses and the Fermi constant. For generality we leave the $\Delta r$-shift component in for discussion however as will be discussed in section 7.1.2 this is shown to be negligible in the parameter spaces considered here and so for practical purposes is omitted from the actual calculation.

### 7.1.1 One-loop Contribution

The expression for the one-loop contribution to $a_\mu$ is given by [55, 56, 57],

$$a_\mu^{\text{2HDM, 1-loop}} = \sum_{\phi=\{h,H,A,H^\pm\}} (Y_l^\phi)^2 \frac{m_\mu}{M_\phi^2} F_\phi(\frac{m_\mu^2}{M_\phi^2}), \tag{7.14}$$

where $\phi$ is an index specifying which Higgs boson is being considered. The summation of $\phi$ in equation (7.14) is relatively straightforward to compute however one must use the correct loop integral for each Higgs. The expressions for these integrals are given by [51],

$$F_{h,H}(x) = \int_0^1 \frac{u^2(2-u)}{1-u+xu^2} du \tag{7.15}$$

$$F_A(x) = \int_0^1 \frac{-u^3}{1-u+xu^2} du \tag{7.16}$$

$$F_{H^\pm}(x) = \int_0^1 \frac{-u(1-u)}{1-(1-u)x} du. \tag{7.17}$$

An important task with the framework was to ensure correct results were being produced, this was achieved throughout the work in a variety of different manners with one example being to check that the one loop result yielded the same results using the general formula in equation (7.14) and the approximated formula (in the valid parameter space with the relevant approximations). The one loop contribution to $a_\mu$ given in equation (7.14) can be approximated as [58],

$$a_\mu^{\text{2HDM, 1-loop}} \approx \left(\frac{\zeta_l}{100}\right)^2 \left(\frac{3.3 + 0.5\ln(\hat{x}_H)}{\hat{x}_H^2} - \frac{3.1 + 0.5\ln(\hat{x}_A)}{\hat{x}_A^2} - \frac{0.04}{\hat{x}_{H^\pm}^2}\right) \times 10^{-10}, \tag{7.18}$$

where $\hat{x}_\phi = \frac{m_\phi}{100\text{ GeV}}$. This approximation of the one loop result served as a valuable method of checking the implementation of the formula into Python was done correctly.

### 7.1.2 $a_\mu^{\Delta\text{r-shift}}$

Although $a_\mu^{\Delta\text{r-shift}}$ component is neglected in our analysis it is important to discuss why this it is treated as such. To begin with one should first examine the expression to calculate it [51],

$$a_\mu^{\Delta r\text{-shift}} = a_\mu^{\text{EW(1)}} \times (-\Delta r^{\text{2HDM}}). \tag{7.19}$$

where $\Delta r^{2HDM}$ is the 2HDM contribution to the fundamental parameter relating electroweak gauge boson masses and the Fermi constant. Ref. [59] shows that $\Delta r^{2HDM}$ is at maximum of order $10^{-3}$. One can calculate the remaining contribution, $a_\mu^{\text{EW(1)}}$ with,

$$a_\mu^{\text{EW(1)}} = \frac{G_F\, m_\mu^2}{8\sqrt{2}\,\pi^2}\Big(\frac{5}{3} + \frac{1}{3}(1 - 4\sin^2(\theta_w))^2\Big). \tag{7.20}$$

which shows $|a_\mu^{\Delta\text{r-shift}}| \leq 2 \times 10^{-12}$ and is hence orders of magnitude smaller than the other contributions to $a_\mu$ so is therefore neglected.

### 7.1.3 Bosonic Contribution

The bosonic contribution to $a_\mu$ in the 2HDM can be split up into the parts calculated from diagrams including the Standard Model Higgs and those containing the additional Higgs bosons, $H, A, H^\pm$. The latter can be further divided into Yukawa dependent and independent components,

$$a_\mu^{\text{B}} = a_\mu^{\text{EW add.}} + a_\mu^{\text{non-Yuk}} + a_\mu^{\text{Yuk}}, \tag{7.21}$$

**Electroweak**

The additional electroweak contribution to $a_\mu$ denoted by $a_\mu^{\text{EW add.}}$ is defined as,

$$a_\mu^{\text{EW add.}} = 2.3 \times 10^{-11}\, \eta\, \zeta_l. \tag{7.22}$$

with $\eta$ defined by equation (7.6) and $\zeta_l$ is taken from table 7.1. This expression for the contribution is given by Feynman diagrams containing gauge/Goldstone/$h$ bosons [51] and as such is described as *Standard Model like*.

It is important to note that although this is a Standard Model like component to the calculation which depends explicitly on the deviation $\eta$ from the Standard Model value, this does not necessarily mean that this contribution will be small relative to the other contributions to $a_\mu^{\text{2HDM}}$ since $\zeta_l$ can enhance the value by a couple orders of magnitude depending on the type of 2HDM considered. Although the opposite effect is also possible for example in a model such as a type-I 2HDM where $\zeta_l = \cot(\beta)$.

**Yukawa-dependent contribution**

The expression for the Yukawa-dependent contribution is obtained by examining the triple Higgs coupling constant, $C_{h,H^\pm,H^\pm}$ in the 2HDM defined in Ref. [58] and observing the fact that in the relevant Feynman diagrams, contributions involving this triple Higgs coupling constant are multiplied by the Yukawa coupling $Y_l^h$ while the other triple Higgs coupling, $C_{H,H^\pm,H^\pm}$ has factors of $Y_l^H$, and $Y_l^A$ attached to it. Using this and the fact that there is a $\tan(\beta)$ enhancement in the triple Higgs coupling constants of the form $(\tan(\beta) - \cot(\beta))$ one arrives at the expression [51],

$$\begin{aligned} a_\mu^{\text{Yuk}} =& a_{0,0}^0 + a_{0,z}^0 \zeta_l \Big( \tan\beta - \cot\beta \Big) + a_{5,0}^0 \Lambda_5 + a_{5,z}^0 \Lambda_5 \zeta_l \big( \tan\beta - \cot\beta \big) + \\ &+ \eta \big( a_{0,0}^2 \big( \tan\beta - \cot\beta \big) + a_{0,z}^1 \zeta_l + a_{5,0}^1 \Lambda_5 \big( \tan\beta - \cot\beta \big) + a_{5,z}^1 \Lambda_5 \zeta_l \big). \end{aligned} \tag{7.23}$$

The $a$ terms above follow notation consistent with that in Ref. [51] which allows for easy identification of the dependencies of these terms by inspection (see appendix I for their analytic expressions). Terms with a superscript $^1$ have a linear dependence on the deviation $\eta$ while those with $^0$ superscript are

independent of $\eta$. The subscript $_z$ denotes that the term is enhanced by $\zeta_l$ and those with a subscript $_5$ are dependent on $\Lambda_5$ which is defined as [51],

$$\Lambda_5 = \frac{2m_{12}^2}{v^2 \sin(\beta)\cos(\beta)}. \tag{7.24}$$

Using equation 7.7 one can rewrite $\Lambda_5$ as a function of the sampled parameters rather than of $m_{12}$ enabling the calculation of $a_\mu^{\text{Yuk}}$,

$$\Lambda_5 = \frac{2}{v^2 \sin(\beta)\cos(\beta)} \left( \frac{M_H^2 \cos(\alpha) - M_h^2 \sin(\alpha) - \lambda_1 v^2 \cos(\beta)}{\tan(\beta)} \right) \tag{7.25}$$

which is the form of the expression used as in this form $\Lambda_5$ is directly calculable from the parameters sampled over in our analysis.

### Yukawa-independent contribution

As the name implies, the Yukawa-independent contribution refers to the part calculated from Feynman diagrams that contain at least one of the new Higgs bosons but without Yukawa couplings. The analytic form for this contribution is given below from Ref. [51],

$$\begin{aligned}
a_\mu^{\text{non-Yuk}} =& \frac{\alpha_{em}^2}{576\,\pi^2\, c_{\text{W}}^4 s_{\text{W}}^4} \frac{m_\mu^2}{M_Z^2} \left\{ \left( \frac{x_A - x_H}{x_A - x_{H^\pm}} \right) \mathcal{T}_2^+(x_A, x_H) + \mathcal{T}_2^-(x_H, x_{H^\pm}) \right. \\
&+ \left( \frac{x_A - x_H}{x_A - x_{H^\pm}} \right) \mathcal{T}_4(x_A, x_{H^\pm}) + \mathcal{T}_4(x_H, x_A) + \mathcal{T}_5(x_{H^\pm}, x_H) + \mathcal{T}_5(x_{H^\pm}, x_A) \\
&+ \mathcal{T}_2^+(x_{H^\pm}, x_H) + \mathcal{T}_2^+(x_{H^\pm}, x_A) + \mathcal{T}_6(x_A, x_{H^\pm}) + \mathcal{T}_6(x_H, x_{H^\pm}) \\
&+ \mathcal{T}_7(x_A, x_H) + \mathcal{T}_7(x_{H^\pm}, x_{H^\pm})(1 - 2c_{\text{W}}^2)^2 + \mathcal{T}_8(x_A, x_{H^\pm}) + \mathcal{T}_8(x_H, x_{H^\pm}) \\
&- \frac{16}{3} c_{\text{W}}^2 s_{\text{W}}^2 (1 + 8c_{\text{W}}^2 - 8c_{\text{W}}^4) + \frac{8 c_{\text{W}}^4 s_{\text{W}}^4}{5 x_{H^\pm}} + f_2 x_{H^\pm} - f_3 x_{H^\pm}^2 \\
&+ f_1(x_A^2 + x_H^2) + f_3 x_{H^\pm}(x_A + x_H) + f_4(x_A + x_H) - f_5 x_A x_H \\
&\left. + \mathcal{T}_1(x_A, x_{H^\pm}) + \mathcal{T}_1(x_H, x_{H^\pm}) + \mathcal{T}_0(x_A, x_{H^\pm}) + \mathcal{T}_0(x_H, x_{H^\pm}) \right\}.
\end{aligned} \tag{7.26}$$

As one can see from above, this contribution is the only bosonic component which contains a dependency on $M_A$. The functions used above in the equation are defined in Appendix H. $s_w$ and $c_w$ denote the sine and cosine of the Weinberg angle respectively. In this project the Weinberg angle was defined

in the standard manner as,

$$\theta_w = \arccos(\frac{M_W}{M_Z}) \tag{7.27}$$

where $M_W$ and $M_Z$ are the mass of the W and Z boson respectively.

### 7.1.4 Fermionic Component

The relevant diagrams to the fermionic contribution contain a single scalar boson interacting with an incoming/outgoing muon and a fermion in the inner loop. This gives a resulting expression proportional to the $Y_f^{\mathcal{S}} Y_l^{\mathcal{S}}$ Yukawa couplings.

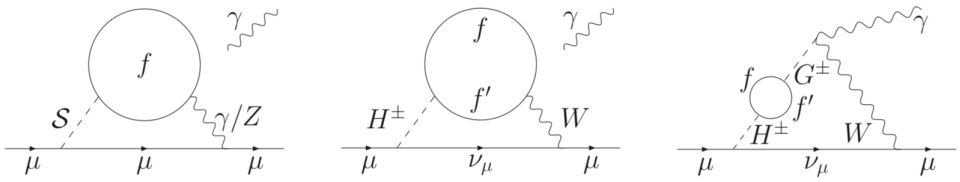



**Figure 7.2:** *Feynman diagrams from [51] for the fermionic component*

Analysis done previously in [54, 60, 61] produced results for the fermionic contribution that are in agreement with those obtained using the expressions in [51].

#### Neutral Higgs Fermionic Contribution

For the analysis of the neutral Higgs component to the fermionic contribution the analytic expression from Ref. [51] is used and reads,

$$\begin{aligned} a_\mu^{\mathrm{F,\,N}} &= \sum_{\mathcal{S}=\{h,H,A\}} \sum_{f=\{u,d,l\}} \left[ f_f^{\mathcal{S}}(M_{\mathcal{S}}, m_f) \right] Y_f^{\mathcal{S}} Y_l^{\mathcal{S}} \\ &\equiv \sum_{\mathcal{S}=\{h,H,A\}} \sum_{f=\{u,d,l\}} \left[ f_{\mathcal{S}}^{\gamma}(M_{\mathcal{S}}, m_f) + f_{\mathcal{S}}^{Z}(M_{\mathcal{S}}, m_f) \right] Y_f^{\mathcal{S}} Y_l^{\mathcal{S}}, \end{aligned} \tag{7.28}$$

where a sum over the all the neutral Higgs is taken for each of the fermions. The functions relevant to 7.28 are defined as,

$$f_{\mathcal{S}}^{\gamma}(M_{\mathcal{S}}, m_f) = \frac{\alpha^2 m_\mu^2}{4\pi^2 M_W^2 s_{\rm W}^2} \left(Q_f^2 N_c^f\right) \left(\frac{m_f^2}{M_{\mathcal{S}}^2}\right) \mathcal{F}_{\mathcal{S}}(M_{\mathcal{S}}, m_f), \tag{7.29}$$

$$f_{\mathcal{S}}^{Z}(M_{\mathcal{S}}, m_f) = \frac{\alpha^2 m_\mu^2}{4\pi^2 M_W^2 s_{\rm W}^2} \left(-\frac{N_c^f Q_f g_v^l g_v^f}{s_{\rm W}^2 c_{\rm W}^2}\right) \times \frac{m_f^2}{(M_{\mathcal{S}}^2 - M_Z^2)} \left[\mathcal{F}_{\mathcal{S}}(M_{\mathcal{S}}, m_f) - \mathcal{F}_{\mathcal{S}}(M_Z, m_f)\right].$$

For $\mathcal{S} = \{h, H\}$ we have

$$\mathcal{F}_{\mathcal{S}}(M_{\mathcal{S}}, m_f) = -2 + \ln\left(\frac{M_{\mathcal{S}}^2}{m_f^2}\right) - \left(\frac{M_{\mathcal{S}}^2 - 2m_f^2}{M_{\mathcal{S}}^2}\right) \frac{\Phi(M_{\mathcal{S}}, m_f, m_f)}{M_{\mathcal{S}}^2 - 4m_f^2}, \tag{7.30}$$

and for $\mathcal{S} = A$

$$\mathcal{F}_{\mathcal{S}}(M_{\mathcal{S}}, m_f) = \frac{\Phi(M_{\mathcal{S}}, m_f, m_f)}{M_{\mathcal{S}}^2 - 4m_f^2}. \tag{7.31}$$

**Charged Higgs Fermion Contribution**

Since there is only a single charged Higgs in the 2HDM considered in this work the only summation required for this part is that over the Standard Model fermions,

$$a_\mu^{\rm F,C} = \sum_{f=\{u,d,l\}} f_f^{H^\pm}(M_{H^\pm}, M_f) Y_f^A Y_l^A, \tag{7.32}$$

where $M_f$ corresponds to pairs of fermions masses,
$M_u = \{(m_u, m_d), (m_c, m_s), (m_t, m_b)\}$, $M_d = M_u$, $M_l = \{(m_e, 0), (m_\mu, 0), (m_\tau, 0)\}$ and

$$f_f^{H^\pm}(M_{H^\pm}, M_f) = \frac{\alpha^2 m_\mu^2}{32\pi^2 M_W^2 s_{\rm W}^4} \frac{N_c^f m_f^2}{(M_{H^\pm}^2 - M_W^2)} \left[\mathcal{F}_f^{H^\pm}(M_f) - \mathcal{F}_f^{W}(M_f)\right], \tag{7.33}$$

where

$$\begin{aligned}
\mathcal{F}_l^{H^\pm}(M_l) &= x_l + x_l\,(x_l-1)\left[\mathrm{Li}_2(1-1/x_l) - \frac{\pi^2}{6}\right] + \left(x_l - \frac{1}{2}\right)\ln(x_l),\\
\mathcal{F}_d^{H^\pm}(M_d) &= -(x_u - x_d) + \left[\frac{\bar{c}}{y} - c\left(\frac{x_u - x_d}{y}\right)\right]\Phi(x_d^{1/2}, x_u^{1/2}, 1)\\
&\quad + c\left[\mathrm{Li}_2\left(1 - \frac{x_d}{x_u}\right) - \frac{1}{2}\ln(x_u)\ln\left(\frac{x_d}{x_u}\right)\right]\\
&\quad + (s + x_d)\ln(x_d) + (s - x_u)\ln(x_u),\\
\mathcal{F}_u^{H^\pm}(M_u) &= \mathcal{F}_d^{H^\pm}(x_u, x_d)\,(Q_u \to 2 + Q_u, Q_d \to 2 + Q_d)\\
&\quad - \frac{4}{3}\left(\frac{x_u - x_d - 1}{y}\right)\Phi(x_d^{1/2}, x_u^{1/2}, 1)\\
&\quad - \frac{1}{3}\left[\ln^2(x_d) - \ln^2(x_u)\right].\\
\mathcal{F}_f^{W}(M_f) &= \mathcal{F}_f^{H^\pm}(M_f)(M_{H^\pm} \to M_W)
\end{aligned}$$

and

$$\begin{aligned}
&x_f \equiv \frac{m_f^2}{M_{H^\pm}^2}, \quad y \equiv (x_u - x_d)^2 - 2\,(x_u + x_d) + 1, \quad s \equiv \frac{(Q_u + Q_d)}{4}\\
&c \equiv \left[(x_u - x_d)^2 - Q_u x_u + Q_d x_d\right], \quad \bar{c} \equiv \left[(x_u - Q_u)\,x_u - (x_d + Q_d)\,x_d\right],\\
&Q_u = \frac{2}{3}, \quad Q_d = -\frac{1}{3},
\end{aligned}$$

Here $f, Q_f$, and $N_f$ follow the standard notation as the fermion, charge of the fermion and colour factor respectively. The definition $g_v^f = \frac{T_3}{2} - Q_f \sin(\theta_w)^2$ is taken from [51] and $\Phi(M_\phi, m_f, m_f)$ is given in appendix H.

Summing over the charged and neutral Higgs contributions and subtracting the contribution of the Standard Model Higgs we arrive at the full two-loop fermionic contribution in the 2HDM,

$$\begin{aligned}
a_\mu^{\mathrm{F}} = \sum_{f=\{u,d,l\}} \Bigg[ \sum_{\mathcal{S}=\{h,H,A\}} & f_f^{\mathcal{S}}(M_{\mathcal{S}}, m_f) Y_f^{\mathcal{S}} Y_l^{\mathcal{S}} + f_f^{H^\pm}(M_{H^\pm}, M_f) Y_f^A Y_l^A\\
& - f_f^{h_{\mathrm{SM}}}(M_{h_{\mathrm{SM}}}, m_f)\Bigg]. \qquad (7.34)
\end{aligned}$$

There will be nine $f_f^{H^\pm}(M_{H^\pm}, M_f)$ terms, three for the up-type quarks, three

for the down-type and three for the leptons, the way this was coded in was to break up the fermion mass pairs into lists,

$$M_u = [m_u, m_c, m_t]$$
$$M_d = [m_d, m_s, m_b]$$
$$M_l = [m_e, m_\mu, m_\tau]$$

So for instance, $\mathcal{F}_d^{H^\pm}(M_d)$ would first be calculated using the first mass from $M_u$ and $M_d$ then added to the contribution using the second mass pair and then the third to give the total.

## 7.2 Lepton Flavour Violation in the Two-Higgs Doublet Model

In a type-III 2HDM each Higgs doublet couples to all fermions and therefore gives rise to flavour violating interactions forbidden in the Standard Model. An important LFV observable to consider in the 2HDM is the Standard Model Higgs decay $h \to \mu\tau$ which CMS reported an excess of the branching ratio which the Standard Model cannot accommodate. This decay can also induce an additional contribution to $a_\mu$ via flavour violating Yukawa couplings, $\rho_e^{\mu\tau}$ and $\rho_e^{\tau\mu}$. These couplings can also have a sizeable effect on the LFV observable $Br(\tau \to \mu\gamma)$ as well as generating other the flavour $\mu - \tau$ violating processes $\tau \to \mu ee$ and $\tau \to 3\mu$. These decays are important to consider alongside $a_\mu$ as the aim of this chapter is to consider a 2HDM which can accommodate both $a_\mu$ and these LFV observables and investigate what the parameter space for such a model looks like so that one is then able to compare this with the MSSM to determine which is more favoured.

### 7.2.1 $Br(h \to \mu\tau)$

The $h \to \mu\tau$ LFV decay is an important process to consider in the 2HDM. Although this decay was not considered in our analysis of the MSSM it is so in this model because the values of the Yukawa couplings $\rho_e^{\mu\tau}$ and $\rho_e^{\tau\mu}$ necessary for the calculation of the LFV decays, are extracted using the equation for the branching ratio,

$$Br(h \to \mu\tau) = \frac{\Gamma(h \to \mu^+\tau^-) + \Gamma(h \to \mu^-\tau^+)}{\Gamma_h} = \frac{\cos(\beta-\alpha)^2 m_h(|\rho_e^{\mu\tau}|^2 + |\rho_e^{\tau\mu}|^2)}{16\pi\Gamma_h} \tag{7.35}$$

using the CMS reported value $Br(h \to \mu\tau) = 0.84\%$ [64], given the restrictions on the value $\cos(\beta-\alpha)$ can take.

### 7.2.2 $Br(\mu \to e\gamma)$

As with supersymmetry, in the 2HDM lepton flavour can be violated whilst keeping the model renormalizable. Since the aim of the project is to build a framework to compare models by calculating the Bayes factor, the observables considered were calculated in both models. The branching ratio of $\mu \to e\gamma$ in the 2HDM is defined as [48],

$$Br(\mu \to e\gamma) = \frac{48\pi^3\alpha_{em}}{G_F^2}(|A_L|^2 + |A_R|^2), \tag{7.36}$$

where $G_f$ is the Fermi constant and $\alpha_{em}$ is the fine-structure constant with the value 1/137 in our analysis. The one-loop contributions to $A_{L,R}$ for $\phi = (h, H, A)$ are given by,

$$A_L^\phi = \frac{1}{16\pi^2}\sum_{i=\mu,\tau}\frac{y^{e*}_{\phi\, ie}}{m_\phi^2}\left[\frac{m_i}{m_\mu}y^{e*}_{\phi\,\mu i}\left(\log\frac{m_\phi^2}{m_i^2} - \frac{3}{2}\right) + \frac{y^e_{\phi\, i\mu}}{6}\right], \tag{7.37}$$

$$A_R^\phi = \frac{1}{16\pi^2}\sum_{i=\mu,\tau}\frac{y^{e}_{\phi\, ei}}{m_\phi^2}\left[\frac{m_i}{m_\mu}y^{e}_{\phi\, i\mu}\left(\log\frac{m_\phi^2}{m_i^2} - \frac{3}{2}\right) + \frac{y^{e*}_{\phi\,\mu i}}{6}\right] \tag{7.38}$$

and for $\phi = H^\pm$ the contribution is,

$$A_L^{H^-} = -\frac{(\bar{\rho}_e^\dagger\rho_e)_{e\mu}}{192\pi^2 m_{H^-}^2}, \quad A_R^{H^-} = 0. \tag{7.39}$$

The Yukawa couplings, $y^f_{\phi ij}$ are given by

$$
\begin{aligned}
y^f_{h\,ij} &= \frac{m^i_f}{v}\sin(\beta-\alpha)\delta_{ij} + \frac{\rho^{ij}_f}{\sqrt{2}}/cos(\beta-\alpha),\\
y^f_{H\,ij} &= \frac{m^i_f}{v}\cos(\beta-\alpha)\delta_{ij} - \frac{\rho^{ij}_f}{\sqrt{2}}\sin(\beta-\alpha),\\
y^f_{A\,ij} &= \begin{cases} -\frac{i\rho^{ij}_f}{\sqrt{2}} \text{ (for } f=u), \\ \frac{i\rho^{ij}_f}{\sqrt{2}} \text{ (for } f=d,\ e) \end{cases}
\end{aligned}
\tag{7.40}
$$

For the parameter spaces concerned in our analysis the two-loop contributions are dominant over the one-loop. The Barr-Zee diagrams were analysed and the two loop component for the $\mu \to e\gamma$ is calculated using the expression [62],

$$
\begin{aligned}
A^{\rm BZ}_L = &-\sum_{\phi=h,H,A;f=u,d,e}\frac{N_C Q_f \alpha}{8\pi^3}\frac{y^{e*}_{\phi\,\mu e}}{m_\mu m_{f_3}}\Big[Q_f\left\{\mathrm{Re}(y^f_{\phi\,33})F_H\left(x_{f\phi}\right) - i\mathrm{Im}(y^f_{\phi\,33})F_A\left(x_{f\phi}\right)\right\}\\
&+\frac{(1-4s^2_W)(2T_{3f}-4Q_f s^2_W)}{16 s^2_W c^2_W}\left\{\mathrm{Re}(y^f_{\phi\,33})\tilde{F}_H\left(x_{f\phi},x_{fZ}\right) - i\mathrm{Im}(y^f_{\phi\,33})\tilde{F}_A\left(x_{f\phi},x_{fZ}\right)\right\}\Big]\\
&+\sum_{\phi=h,H}\frac{\alpha}{16\pi^3}\frac{g_{\phi WW}y^{e*}_{\phi\,\mu e}}{m_\mu v}\Big[3F_H\left(x_{W\phi}\right)+\frac{23}{4}F_A\left(x_{W\phi}\right)\\
&\qquad\qquad +\frac{3}{4}G\left(x_{W\phi}\right)+\frac{m^2_\phi}{2m^2_W}\left\{F_H\left(x_{W\phi}\right)-F_A\left(x_{W\phi}\right)\right\}\\
&+\frac{1-4s^2_W}{8s^2_W}\left\{\left(5-t^2_W+\frac{1-t^2_W}{2x_{W\phi}}\right)\tilde{F}_H(x_{W\phi},x_{WZ})\right.\\
&\left.+\left(7-3t^2_W-\frac{1-t^2_W}{2x_{W\phi}}\right)\tilde{F}_A(x_{W\phi},x_{WZ})+\frac{3}{2}\left\{F_A(x_{W\phi})+G(x_{W\phi})\right\}\right\}\Big],
\end{aligned}
\tag{7.41}
$$

$$
A^{\rm BZ}_R = A^{\rm BZ}_L(y^{e*}_{\phi\,\tau\mu}\to y^e_{\phi\,\mu\tau},\ i\to -i),
\tag{7.42}
$$

where $x_{f\phi}=m^2_{f_3}/m^2_\phi$, $x_{fZ}=m^2_{f_3}/m^2_Z$ ($f_3=t,b,\tau$ for $f=u,d,e$), $x_{W\phi}=m^2_W/m^2_\phi$ and $x_{WZ}=m^2_W/m^2_Z$, and $s^2_W=\sin\theta^2_W$, $c^2_W=\cos\theta^2_W$ and $t^2_W=\tan\theta^2_W$. $T_{3f}$ denotes the isospin of the fermion. The couplings $g_{\phi WW}=s_{\beta\alpha}$ ($c_{\beta\alpha}$) for $\phi=h$ ($\phi=H$).

The loop functions used to calculate the branching ratio in equation (7.41),

$F_{H,\ A}$, $G$ and $\tilde{F}_{H,\ A}$, are integrals defined as,

$$F_H(z) = \frac{z}{2}\int_0^1 dx \frac{1-2x(1-x)}{x(1-x)-z}\log\frac{x(1-x)}{z}, \tag{7.43}$$

$$F_A(z) = \frac{z}{2}\int_0^1 dx \frac{1}{x(1-x)-z}\log\frac{x(1-x)}{z}, \tag{7.44}$$

$$G(z) = -\frac{z}{2}\int_0^1 dx \frac{1}{x(1-x)-z}\left[1-\frac{z}{x(1-x)-z}\log\frac{x(1-x)}{z}\right], \tag{7.45}$$

$$\tilde{F}_H(x,y) = \frac{xF_H(y)-yF_H(x)}{x-y}, \tag{7.46}$$

$$\tilde{F}_A(x,y) = \frac{xF_A(y)-yF_A(x)}{x-y}. \tag{7.47}$$

The total $A_{L,\ R}$ is the sum of all contributions;

$$A_{L,\ R} = \sum_{\phi=h,H,A,H^{\pm}} A^{\phi}_{L,\ R} + A^{\mathrm{BZ}}_{L,\ R}. \tag{7.48}$$

which then allows one to calculate the branching ratio for $\mu \to e\gamma$ by using equation (7.36).

### 7.2.3 $Br(\tau \to \mu\gamma)$

Since there is no analytic expression for these off diagonal couplings they are treated as free parameters in our analysis. Constraints on these Yukawa couplings were obtained by examining different lepton flavour violating observables individually that use these couplings and applying known experimental constraints on the observable values to constrain the allowed values of the couplings. Although MEG has set the most stringent limit on a lepton flavour violating process ($\mu \to e\gamma$) there are experiments planned in the near future to attempt to improve on the measured limits of LFV decays in the $\tau-$ channel [65]. One of the $\tau-$LFV observables considered was the branching ratio $\tau \to \mu\gamma$ given by,

$$\frac{Br(\tau \to \mu\gamma)}{Br(\tau \to \mu\bar{\nu}_\mu\nu_\tau)} = \frac{48\pi^3\alpha\left(|A_L|^2+|A_R|^2\right)}{G_F^2}. \tag{7.49}$$

The one loop contributions to $A_{L,R}$ are given by [76]

$$
\begin{aligned}
A_{L,\ R} &= \sum_{\phi=h,\ H,\ A,\ H^-} A^{\phi}_{L,\ R}, \qquad (7.50)\\
A^{\phi}_{L} &= \frac{y^{e*}_{\phi\ \tau\mu}}{16\pi^2 m^2_{\phi}}\left[y^{e*}_{\phi\ \tau\tau}\left(\log\frac{m^2_{\phi}}{m^2_{\tau}}-\frac{3}{2}\right)+\frac{y^{e}_{\phi\ \tau\tau}}{6}\right], \quad (\phi=h,\ H,\ A)\\
A^{\phi}_{R} &= A^{\phi}_{L}|_{y^{e*}_{\phi\ \tau\mu}\to y^{e}_{\phi\ \mu\tau},\ \ y^{e}_{\phi\ \tau\tau}\leftrightarrow y^{e*}_{\phi\ \tau\tau}}, \quad (\phi=h,\ H,\ A),\\
A^{H^-}_{L} &= -\frac{(\rho^{\dagger}_{e}\rho_{e})^{\mu\tau}}{192\pi^2 m^2_{H^-}}, \quad A^{H^-}_{R}=0.
\end{aligned}
$$

The two-loop contributions to $A_L L, R$ are the same as for $\mu\to e\gamma$ given by equation (7.41) except $y^{e(*)}_{\phi\ \mu e\ (e\mu)}$ is replaced by $y^{e(*)}_{\phi\ \tau\mu\ (\mu\tau)}$, and the $\mu$ mass ($m_\mu$) should be replaced by the $\tau$ mass ($m_\tau$).

### 7.2.4 $\tau\to 3\mu$ and $\tau\to\mu ee$

The experimental study of $\tau\to lll$ was started with MARKII [73] these types of decays have previously been discussed in [66, 67] for the 2HDM.

From [77] we obtain the following forms for the branching ratios for the remaining two $\tau$ decays considered in the MSSM. The branching ratio for lepton flavour violating $\tau-$decays is divided by the standard $\tau$ decay to give the decay fraction,

$$
\begin{aligned}
\frac{\mathrm{Br}(\tau\to 3\mu)}{\mathrm{BR}(\tau\to\mu\nu\bar{\nu})} &= \sum_{\phi,\ \phi'=h,\ H,\ A}\frac{I(\phi,\ \phi')}{64G^2_F},\\
I(\phi,\ \phi') &= 2\left(\frac{y^{e}_{\phi\mu\tau}y^{e*}_{\phi\mu\mu}}{m^2_{\phi}}\right)\left(\frac{y^{e*}_{\phi'\mu\tau}y^{e}_{\phi'\mu\mu}}{m^2_{\phi'}}\right)+2\left(\frac{y^{e}_{\phi\tau\mu}y^{e*}_{\phi\mu\mu}}{m^2_{\phi}}\right)\left(\frac{y^{e*}_{\phi'\tau\mu}y^{e}_{\phi'\mu\mu}}{m^2_{\phi'}}\right)\\
&\quad+\left(\frac{y^{e}_{\phi\mu\tau}y^{e}_{\phi\mu\mu}}{m^2_{\phi}}\right)\left(\frac{y^{e*}_{\phi'\mu\tau}y^{e*}_{\phi'\mu\mu}}{m^2_{\phi'}}\right)+\left(\frac{y^{e}_{\phi\tau\mu}y^{e}_{\phi\mu\mu}}{m^2_{\phi}}\right)\left(\frac{y^{e*}_{\phi'\tau\mu}y^{e*}_{\phi'\mu\mu}}{m^2_{\phi'}}\right),
\end{aligned}
\qquad (7.51)
$$

$$\frac{\mathrm{Br}(\tau \to \mu ee)}{\mathrm{BR}(\tau \to \mu\nu\bar{\nu})} = \sum_{\phi,\ \phi'=h,\ H,\ A} \frac{J(\phi,\ \phi')}{32G_F^2},$$

$$J(\phi,\ \phi') = \left(\frac{y^e_{\phi\mu\tau} y^{e*}_{\phi ee}}{m^2_\phi}\right)\left(\frac{y^{e*}_{\phi'\mu\tau} y^e_{\phi' ee}}{m^2_{\phi'}}\right) + \left(\frac{y^e_{\phi\tau\mu} y^{e*}_{\phi ee}}{m^2_\phi}\right)\left(\frac{y^{e*}_{\phi'\tau\mu} y^e_{\phi' ee}}{m^2_{\phi'}}\right)$$
$$+ \left(\frac{y^e_{\phi\mu\tau} y^e_{\phi ee}}{m^2_\phi}\right)\left(\frac{y^{e*}_{\phi'\mu\tau} y^{e*}_{\phi' ee}}{m^2_{\phi'}}\right) + \left(\frac{y^e_{\phi\tau\mu} y^e_{\phi ee}}{m^2_\phi}\right)\left(\frac{y^{e*}_{\phi'\tau\mu} y^{e*}_{\phi' ee}}{m^2_{\phi'}}\right). \tag{7.52}$$

The above LFV decays $Br(\tau \to 3\mu)$ and $Br(\tau \to \mu ee)$ constrain the Yukawa couplings $\rho_e^{ee}$ and $\rho_e^{\mu\mu}$ by imposing the experimental bounds on these decays.

### 7.2.5 Lepton Flavour Violating Contribution to $a_\mu$

The section began with a brief discussion on LFV in the 2HDM and the possible effects this can have on the model. One such impact mentioned was the ability to give rise to additional contributions to $a_\mu$ induced by the LFV Yukawa couplings. The off diagonal Yukawa couplings introduced to calculate the lepton flavour violating observables discussed earlier give rise to an additional contribution to $a_\mu$ given by the diagram [69],

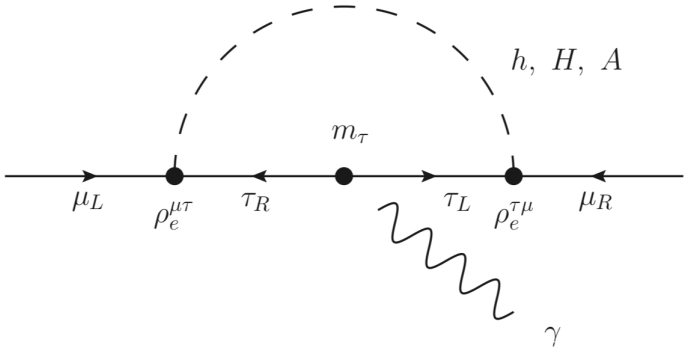



**Figure 7.3:** *Feynman diagrams from [48] for contribution to* $a_\mu$

with the analytical expression as [69],

$$a_\mu^{\mathrm{LFV}} = \frac{m_\mu m_\tau \rho_e^{\mu\tau} \rho_e^{\tau\mu}}{16\pi^2}\left[\frac{c^2_{\beta\alpha}(\log\frac{m_h^2}{m_\tau^2} - \frac{3}{2})}{m_h^2} + \frac{s^2_{\beta\alpha}(\log\frac{m_H^2}{m_\tau^2} - \frac{3}{2})}{m_H^2} - \frac{\log\frac{m_A^2}{m_\tau^2} - \frac{3}{2}}{m_A^2}\right] \tag{7.53}$$

Equation (7.53) shows that the only Yukawa couplings that contribute to the anomalous magnetic moment of the muon contribution from the $\tau \to \mu\gamma$ are those constrained by the Higgs decay, $h \to \mu\tau$. For the excess reported by CMS and ATLAS of the $h \to \mu\tau$ decay the Yukawa couplings must fulfil [69],

$$\begin{aligned} \bar{\rho}^{\mu\tau} &= \sqrt{\frac{|\rho_e^{\mu\tau}|^2 + |\rho_e^{\tau\mu}|^2}{2}}, \\ &\simeq 0.26\left(\frac{0.01}{\cos(\beta-\alpha)}\right)\sqrt{\frac{Br(h\to\mu\tau)}{0.84\times10^{-2}}}, \end{aligned} \tag{7.54}$$

which indicates that the introduction of $\mu-\tau$ flavour violation leads to Yukawa couplings whose contribution could help toward explaining the anomalous magnetic moment of the muon. Note we adopt the convention that $\rho_e^{\mu\tau}\rho_e^{\mu\tau} < 0$ with $\rho_e^{\mu\tau} = -\rho_e^{\mu\tau}$ to maintain $a_\mu^{\text{LFV}} > 0$. Combining equations (7.13) and (7.53) one obtains the expression for $a_\mu$ in the two Higgs doublet model with lepton flavour violation,

$$a_\mu^{2HDM} = a_\mu^{\text{LFV}} + a_\mu^{2HDM,1-loop} + a_\mu^{\text{B}} + a_\mu^{\text{F}} + a_\mu^{\Delta r\text{-shift}}. \tag{7.55}$$

## 7.3 Two-Higgs Doublet Input

The aim of this project is to demonstrate how to compare two different models and show the result of this between the MSSM and 2HDM. Recall the definition of the discrete marginalized likelihood from equation (2.5).

$$P(D|M_i) = \sum_{\vec{\theta}_i} P(D|\vec{\theta}_i)P(\vec{\theta}_i|M_i)$$

here $P(D|\vec{\theta}_i)$ is the target distribution we fit our sampling to and it represents the probability of the data given the parameters in the model. The Bayes factor offers a method of determining which model is favoured given the data meaning that any results or bounds for the observables is the same regardless of the models. This means that the target distributions for the 2HDM are the same for the MSSM presented earlier in section 6.4, so for $a_\mu$ we have $\mu = 26.8\times10^{-10}$ and $\sigma = 7.6\times10^{-10}$ with a distribution,

$$P(a_\mu|\theta_{\text{2HDM}}) = \frac{1}{\sqrt{2\pi(7.6\times10^{-10})^2}} e^{-\frac{(x-26.8\times10^{-10})^2}{2(7.6\times10^{-10})^2}}, \tag{7.56}$$

with $\theta_{\text{2HDM}}$ denoting the 2HDM parameters and $x$ representing the calculated value in the 2HDM in our analysis.

The following range of parameters were chosen in line with existing constraints. The Large Electron-Positron Collider (LEP) [17] has given a model-independent constraint on the charged scalar mass from the charged Higgs decays, $H^{\pm} \to \tau\nu$ and $H^{\pm} \to cs$ [71],

$$M_{H^{\pm}} \geq 80\text{GeV}. \tag{7.57}$$

The couplings are constrained by perturbativity which requires that they are below a maximum value of,

$$|\lambda_i| < 4\pi \quad [78]. \tag{7.58}$$

The masses of the non Standard Model Higgs bosons are sampled with an upper bound of 1 TeV to give as broad a sample space as possible without leading to these heavy scalars decoupling from the Standard Model sector [68].

The mixing angle $\beta - \alpha$ is indicated to be small by LHC measurements of Higgs couplings and flavour violating searches. For our analysis $\cos(\beta - \alpha)$ is taken to be of order $|\cos(\beta - \alpha)| \lesssim \mathcal{O}(0.01)$ which results in the coupling $\bar{\rho}^{\mu\tau}$ calculated in equation 7.54 being of order $\mathcal{O}(1)$ which fits the observed CMS observed excess in $h \to \mu\tau$ [69]. If $\cos(\beta - \alpha)$ is taken to have a larger interval of allowed values then this results in parameter spaces where experimental constraints on observables are violated.

The lower limit on the allowed mass for $M_H$ is 125 GeV is because we denote the lightest neutral Higgs boson in the 2HDM model as being the standard model Higgs, $h = h_{\text{SM}}$ so by definition then $M_H > M_h$.

The value of $\tan(\beta)$ is initially restricted by the requirement that the top and bottom Yukawa couplings, $y_t$ and $y_b$, are perturbative [70],

$$\frac{y_i^2}{4\pi} \lesssim 1 \tag{7.59}$$

with the form of the Yukawa couplings being,

$$y_t = \frac{\sqrt{2}m_t}{v\sin(\beta)} \tag{7.60}$$

$$y_b = \frac{\sqrt{2}m_b}{v\cos(\beta)} \tag{7.61}$$

here $m_t$ and $m_b$ are the top and bottom mass with $v = 246$. Taking the particle masses as $m_t = 171$ GeV and $m_b = 4.20$ GeV [17] one obtains an initial bound on $\tan(\beta)$ of,

$$0.28 \lesssim \tan(\beta) \lesssim 147. \tag{7.62}$$

Additional constraints on $\tan(\beta)$ come from the alignment parameter $\zeta_l$ which to satisfy LHC data must obey [71],

$$|\zeta_l| < 100,$$

From table 7.1 we see that for a type-III 2HDM that is considered for LFV, $\tan(\beta) = -\zeta_l$ which gives the final range on $\tan(\beta)$ used in the 2HDM for this work as,

$$0.28 \lesssim \tan(\beta) \lesssim 100. \tag{7.63}$$

The off diagonal Yukawa couplings are constrained by the LFV observables they are most dominant for, for instance $\rho_e^{\mu e}$ and $\rho_e^{e\mu}$ are constrained by the experimental limit $Br(\mu \to e\gamma) < 4.2 \times 10^{-13}$ to a bound of,

$$-3 \times 10^{-4} \leq \rho_e^{\mu e}, \rho_e^{e\mu} \leq 3 \times 10^{-4},$$

as well as $\rho_e^{e\tau}$ and $\rho_e^{e\tau}$ constrained by this decay to,

$$-1 \times 10^{-5} \leq \rho_e^{e\tau}, \rho_e^{e\tau} \leq 1 \times 10^{-5}.$$

Similarly $h \to \mu\tau$ constrains the Yukawa couplings $\rho_u^{tt}$ and $\rho_e^{\tau\tau}$ to,

$$-0.2 \leq \rho_u^{tt} \leq 0.2$$
$$-0.4 \leq \rho_u^{\tau\tau} \leq 0.4.$$

Finally, the experimental limit on the observables [49].,

$$Br(\tau \to 3\mu) < 2.1 \times 10^{-8}, \quad Br(\tau \to \mu ee) < 1.8 \times 10^{-8}, \tag{7.64}$$

gives us constraints on $\rho_l^{\mu\mu}$ and $\rho_l^{ee}$ of,

$$0 < \rho_e^{\mu\mu}, \rho_e^{ee} < 0.006.$$

The constraints above on the Yukawa couplings from the experimental bounds are taken from Ref. [48].

| Parameter | Range | Prior, $P(\theta_i\|M)$ |
|---|---|---|
| $M_{H^\pm}$ | $[80, 1000]$ GeV | $\begin{cases} \frac{1}{920} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $M_H$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $M_A$ | $[0, 1000]$ GeV | $\begin{cases} \frac{1}{1000} \ \forall\theta_i \in [0, 1000], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\cos(\beta - \alpha)$ | $[-0.01, 0.01]$ | $\begin{cases} \frac{1}{0.02} \ \forall\theta_i \in [-0.01, 0.01], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\lambda_1$ | $[-4\pi, 4\pi]$ | $\begin{cases} \frac{1}{8\pi} \ \forall\theta_i \in [-4\pi, 4\pi], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\tan(\beta)$ | $[1, 100]$ | $\begin{cases} \frac{1}{100} \ \forall\theta_i \in [0, 100], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{\mu\mu}$ | $[0, 0.006]$ | $\begin{cases} \frac{1}{0.006} \ \forall\theta_i \in [0, 0.06], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{ee}$ | $[0, 0.006]$ | $\begin{cases} \frac{1}{0.006} \ \forall\theta_i \in [0, 0.06], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{tt}$ | $[-0.2, 0.2]$ | $\begin{cases} \frac{1}{0.4} \ \forall\theta_i \in [-0.2, 0.2], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{\tau\tau}$ | $[-0.4, 0.4]$ | $\begin{cases} \frac{1}{0.8} \ \forall\theta_i \in [-0.4, 0.4], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{\mu e}$ | $[-3 \times 10^{-4}, 3 \times 10^{-4}]$ | $\begin{cases} \frac{1}{6\times10^{-4}} \ \forall\theta_i \in [-3 \times 10^{-4}, 3 \times 10^{-4}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{e\mu}$ | $[-3 \times 10^{-4}, 3 \times 10^{-4}]$ | $\begin{cases} \frac{1}{6\times10^{-4}} \ \forall\theta_i \in [-3 \times 10^{-4}, 3 \times 10^{-4}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{e\tau}$ | $[-1 \times 10^{-5}, 1 \times 10^{-5}]$ | $\begin{cases} \frac{1}{2\times10^{-5}} \ \forall\theta_i \in [-1 \times 10^{-5}, 1 \times 10^{-5}], \\ 0 \quad \text{otherwise} \end{cases}$ |
| $\rho_e^{\tau e}$ | $[-1 \times 10^{-5}, 1 \times 10^{-5}]$ | $\begin{cases} \frac{1}{2\times10^{-5}} \ \forall\theta_i \in [-1 \times 10^{-5}, 1 \times 10^{-5}], \\ 0 \quad \text{otherwise} \end{cases}$ |

**Table 7.2:** *Table showing the 2HDM parameters used in our calculations along with the ranges they were sampled over and their prior distribution denoted as $P(\theta_i|M)$ with $\theta_i$ representing the sampled value of the parameter.*

One of the aims of this project is to calculate the Bayes factor between the MSSM and 2HDM. To draw a meaningful comparison between the models they are both given the same data to fit. This data given to the models are the

observables and the experimental bounds/measurements. Since such data is model independent then the target distribution for the 2HDM is the same as for the MSSM given in equation (6.61),

$$P(D|\vec{\theta}_{\text{2HDM}}) = P(a_\mu|\vec{\theta}_{\text{2HDM}})P(Br(\mu \to e\gamma)|\vec{\theta}_{\text{2HDM}})P(Br(\tau \to \mu\gamma)|\vec{\theta}_{\text{2HDM}}) \times P(Br(\tau \to 3\mu)|\vec{\theta}_{\text{2HDM}})P(Br(\tau \to \mu ee)|\vec{\theta}_{\text{2HDM}}). \tag{7.65}$$

As with the MSSM, this is used to examine the parameter space of the model in which both these LFV processes can occur and the anomalous magnetic moment of the muon is explained.

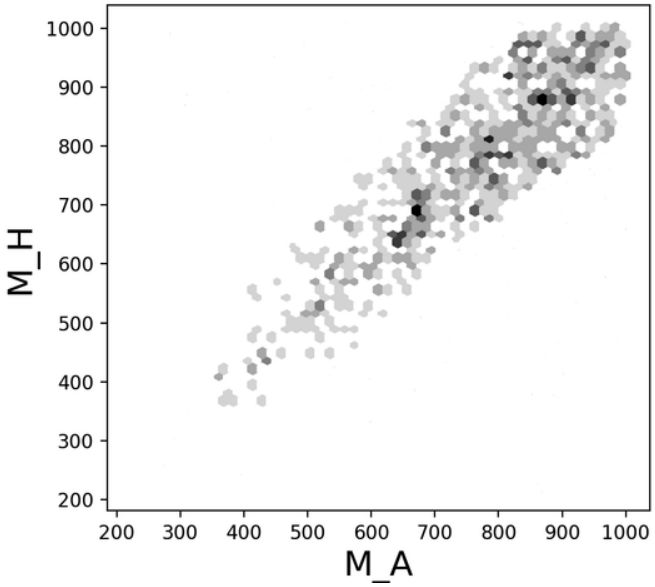



**Figure 7.4:** *Plot showing* $10^5$ *sampled values for the masses* $M_H$ *and* $M_A$ *in GeV for the 2HDM constrained by experimental measurements of* $a_\mu$ *and the bounds on LFV observables.*

A previously discussed, hexbin plots such as those in figures 7.4 and 7.5 are useful for interpreting results as they show which values on the axis are preferred by displaying the point density in each bin which is also correlated to the likelihood. Examining the masses we find no excluded values as such since large values of one mass parameter does not necessarily lead to excluded observable values since the effect on the observables of one mass parameter being large can be offset by other mass parameters having a relatively low value. When examining $M_A$ and $M_H$ together we find that for a 2HDM to fit the observable data given here, that values $M_A \simeq M_H$ as shown in Fig. 7.4 are

preferred in the model which corresponds to low values of $\cos(\beta - \alpha)$ required for LFV.

Fig. 7.5 shows $\cos(\beta - \alpha)$ against $\tan(\beta)$ and one can see that values of $\cos(\beta - \alpha)$ close to zero are absent. This comes from the fact that equation 7.35 contains $\cos(\beta - \alpha)$ in the denominator so as $\cos(\beta - \alpha) \to 0$ then the the Yukawa couplings, $\rho_e^{\{\mu\tau,\tau\mu\}}$ tend to large values and when used in the LFV decays of the $\tau$ lepton in equations (7.49), (7.51) and (7.52) lead to values above the experimental bounds and are hence excluded.

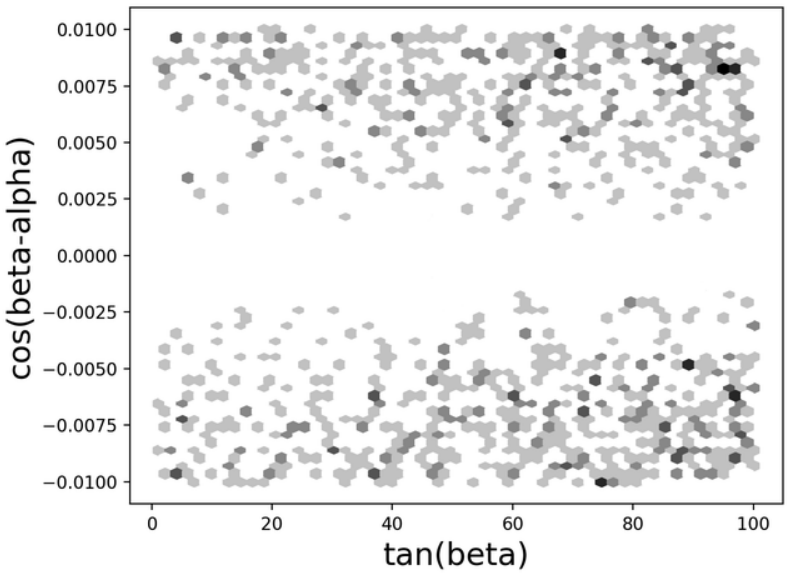



**Figure 7.5:** *Plot showing* $\cos(\beta - \alpha)$ *against* $\tan(\beta)$ *(*$10^5$ *points).*

The coupling $\lambda_1$ is does not acquire any additional restrictions with the introduction of additional LFV observables since it contributes only to $\Lambda_5$ when calculating $a_\mu^{\text{Yuk}}$ in equation (7.23). The off-diagonal Yukawa couplings are not restricted further than discussed in table 7.2 since the bounds used are informed by the LFV decay to which these couplings are dominant in. For instance $\rho_e^{\{\mu e,e\mu\}}$ are the dominant Yukawa couplings in the calculation of $Br(\mu \to e\gamma)$ and so any bounds on the $\mu \to e\gamma$ decay restricts only these since other Yukawa couplings have minimal impact on this calculation and so the introduction of additional LFV decays does not constrain these couplings further since these couplings do not have a significant impact on other decays. After running the code for the 2HDM with the input given in this section one is then able to calculate the Bayes factor between the 2HDM and the MSSM.

# Chapter 8

# Conclusions

The Bayes factor provides a useful method for model comparison by interpreting a simple numerical output from table 2.1. The value of this number indicates which of the two models is preferred and to what degree given the data. Section 2.3 provided a detail discussion on the Bayes factor with the underlying statistics and gave the definition for the Bayes factor as the ratio of marginalized likelihoods defined as,

$$P(D|M_i) = \int_{\vec{\theta}_i} P(D|\vec{\theta}_i)P(\vec{\theta}|M_i)d\vec{\theta}_i. \tag{8.1}$$

Here $D$ represents the data which in our case are the experimental measurements and bounds, $M_i$ denotes the model with $\vec{\theta}_i$ being the corresponding parameter vectors of the model. As discussed previously since this project deals with discrete sample sets then the marginalized likelihood becomes a summation rather than an integral and so the Bayes factor between the two models has the following definition,

$$B_{12} = \frac{\Sigma P(D|\vec{\theta}_1, M_1)P(\vec{\theta}_1|M_1)}{\Sigma P(D|\vec{\theta}_2, M_2)P(\vec{\theta}_2|M_2)} \tag{8.2}$$

Recall the prior probabilities $P(\vec{\theta}_i|M_i)$ in the equation above represent the probability distribution the parameters are sampled according to. The priors for each of the individual parameters are given in tables 6.1 and 7.2 for the MSSM and 2HDM respectively with the total prior for the model calculated as,

$$P(\vec{\theta}_i|M_i) = \prod_{j=1} P(\theta_{1j}|M_i) \tag{8.3}$$

with the index $i$ denoting the model and $j$ denoting the parameter. The total prior for the models are given by the product of the priors for the uncorrelated model parameters. When multiplying the nonzero values in table 7.2 and 6.1 we obtain values for the total priors of,

$$P(\vec{\theta}_{\text{2HDM}}|M_{\text{2HDM}}) = 1.30355 \times 10^{10} \tag{8.4}$$

and

$$P(\vec{\theta}_{\text{MSSM}}|M_{\text{MSSM}}) = 1.51172 \times 10^{-22} \tag{8.5}$$

One can see from the large orders of magnitude difference and the fact that equation (8.4) is above one that some normalization is required. Since the analysis deals with discrete data sets the normalization of the probability is achieved by dividing by the sum of all the prior values of the data set.

$$P'(\vec{\theta}_{\text{2HDM}}|M_{\text{2HDM}}) = \frac{P(\vec{\theta}_{\text{2HDM}}|M_{\text{2HDM}})}{\sum_n P(\vec{\theta}_{\text{2HDM}}|M_{\text{2HDM}})}. \tag{8.6}$$

which gives the normalized prior for the 2HDM as

$$P'(\vec{\theta}_{\text{2HDM}}|M_{\text{2HDM}}) = \frac{1}{n} \tag{8.7}$$

where $n$ is the number of points in the data set. This value of the prior reflects the fact that the probability of one value occurring in the set is the same across the entire uniform distribution with the probability of all events in the space summing to one. Normalizing the prior distribution in this manner also removes the dependency of how the model is parameterised, for example in the 2HDM one can parameterise the model in terms of the masses or the couplings $\lambda_i$ and the unormalized uniform distribution for each of these are orders of magnitude apart. After normalizing the MSSM prior distribution given in equation (8.5) one obtains the same result for the prior as in equation (8.7) which leads to the Bayes factor having the form,

$$B_{12} = \frac{\Sigma P(D|\vec{\theta}_1, M_1)}{\Sigma P(D|\vec{\theta}_2, M_2)} \tag{8.8}$$

Here $P(D|\vec{\theta}_i, M_i)$ represents the target distribution of our sampling in the code and as discussed in sections 6.4 and 7.3 is given by,

$$P(D|\vec{\theta}_i) = P(a_\mu|\vec{\theta}_i)P(Br(\mu \to e\gamma)|\vec{\theta}_i)P(Br(\tau \to \mu\gamma)|\vec{\theta}_i) \times P(Br(\tau \to 3\mu)|\vec{\theta}_i)P(Br(\tau \to \mu ee)|\vec{\theta}_i). \tag{8.9}$$

This cancellation of the uniform priors can be interpreted as all possible parameter values have the same probability after normalization. As pointed out in equation (8.9) the total likelihood for the models in this analysis is the product of the individual likelihoods of the observables. Using this along with the fact that the prior probabilities are uniform distributions which cancel out after normalizing (see equation (8.8)) one can rewrite equation (8.2) as,

$$B_{12} = \prod_{\text{Observables}} \frac{\Sigma P(D|\vec{\theta}_1, M_1)}{\Sigma P(D|\vec{\theta}_2, M_2)} \tag{8.10}$$

meaning the Bayes factor for the model is simply the product of the Bayes factor in the columns for the individual observables in table 8.1. This allows one to determine if any of the observables play the role of a crucial discriminator when comparing the models in this manner. Below is a table with the values of the Bayes factor between the 2HDM and MSSM containing only a single observable with the end column showing the Bayes factor between the models with all the discussed observables incorporated.

| | $a_\mu$ | $Br(\mu \to e\gamma)$ | $Br(\tau \to \mu\gamma)$ | $Br(\tau \to 3\mu)$ | $Br(\tau \to \mu ee)$ | Total |
|---|---|---|---|---|---|---|
| $B_{12}$ | 2.20 | 2.90 | 1.10 | 1.27 | 0.96 | 12.94 |

**Table 8.1:** *Table showing the Bayes factor between the 2HDM and MSSM when considering only one observable with the end column containing the Bayes factor between the models when all the above observables are considered.*

Table 8.1 shows that the total Bayes factor between the 2HDM and MSSM gives the result, interpreted from table 2.1, that the 2HDM has substantial evidence in its favour given the data on these observables. This result could be thought of as a reflection on the fact that parts of the MSSM parameter space presented are heavily restricted by requiring $\mu > 0$ for results of $a_\mu$ to be positive otherwise this result would be strongly disfavoured by reported experimental results.

From table 8.1 one can see the crucial discriminators in the resultant Bayes factor are $a_\mu$ and $Br(\mu \to e\gamma)$. As discussed the $a_\mu$ in the MSSM disfavours values of $\mu < 0$ and as demonstrated in figures 6.1 and 6.2 there is a strong correlation between values of $a_\mu$ and $Br(\mu \to e\gamma)$ meaning that parameter spaces disfavoured by $a_\mu$ can also correspond to parameter spaces disfavoured by $Br(\mu \to e\gamma)$ since higher values of parameters lead to higher values of both observables. If all the MSSM masses were to be in the upper part of the bounds then this would require high values $\tan(\beta)$ to explain $a_\mu$ which would violate the SM hierarchy in equation (6.54) and hence is a disfavoured part of the parameter space.The Bayes factors for the $\tau-$decay observables is approximately equal to one which from table 2.1 tells us that there is very little evidence that the 2HDM is favoured when considering only $Br(\tau \to \mu\gamma)$ and $Br(\tau \to 3\mu)$ (and negligible evidence that the MSSM is favoured when considering $Br(\tau \to \mu ee)$). This reflects that neither the 2HDM or MSSM parameter spaces presented have any large portions excluded when calculating these observables.

The number of parameters a model has plays a significant role in statistical comparison of this nature. In general a model with good theoretical justification that is capable of fitting the data with fewer parameters would be preferred over a model requiring more parameters to fit the data. One can view this as indicating a possible deficiency of one model over another if one model required the introduction of additional parameters to explain the same data as the model that required fewer. This fact was echoed in the results of the thesis reflecting that the 2HDM, with fewer parameters, was favoured over the MSSM when fitting the same data. It is important to remember that the Bayes factor only reflects the degree to which one model is favoured over another, it does not contain the information as to which is the correct model. It could indeed be the case that one point in the MSSM parameter space represents the values measured in the future that are determined to be the correct parameter values for explaining the observables. However the statistical analysis in this work does not offer such insight, only that the evidence suggests to the contrary that the 2HDM is strongly favoured given the current data.

The constraints employed on the parameters were chosen to be as broad as possible as to prevent any bias being introduced. Since these parameters have no measured value the only information available was to use constraints from

experimental and theoretical work. The impact of this was to give a model paramterised in the least biased way possible whilst also retaining important physical constraints, for example not having the masses too high as to cause the theory to decouple. Parameter ranges chosen to be smaller would be arbitrary as there was no information to indicate this was an appropriate approach and so would lead to bias in the result of the Bayes factor. Conversely larger ranges of parameter values would lead to issues with the underlying physics of the model or disagreement with theoretical and experimental constraints on the parameters.

The motivation behind considering the 2HDM and MSSM was their ability to explain certain deficiencies of the Standard Model. Both of these models offered solutions to the anomalous magnetic moment of the muon and the observed non-zero values for the branching ratio of lepton flavour violating decays. If one were to calculate the Bayes factor between either of these two models and the Standard Model one would find, through similar arguments presented in section 5.7 the same result as equation (5.23), that the MSSM and 2HDM are decisively favoured over the Standard Model when trying to explain $a_\mu$ and the LFV decays.

The computational framework presented here was designed to be as general as possible to allow a user to consider any physical model with any number of observables calculated with any number of parameters. This approach taken when constructing the framework means that the user is able to calculate the Bayes factor between any two models, not just the 2HDM and MSSM. This is a crucial aspect of this framework given the large amount of current and future planned experimental work since statistical analysis of this nature is an important task to determine which BSM theories are favoured by data and hence better candidates for explaining the deficiencies of the Standard Model.

# Appendix A

# Imports

The first thing to do in the code is import the packages required in the code. Below we outline the packages used in the project and give an overview of the role they play.

- **NumPy** - A scientific package that gives the ability to construct matrices with the $np.array([[...]])$ command. Operations such as transposing and matrix multiplication are also available from this package. Similarly the package allowed operations on vectors or lists of numbers, $np.sum[...]$ allowed the sum of all elements in the list and $np.prod[...]$ produced the value of multiplying all the elements together.

- **pypmc** - This is the package that handled the sampling part of our project. The tools for performing the Markov-Chain Monte Carlo sampling were taken from pypmc.

- **scipy** - This package contains some of the tools for the mathematics such as the integration.

- **math** - The math package contains mathematical functions such as trigonometric and logarithmic ones. It also allows one to perform mathematical operations such as taking the square root.

- **cmath** - A package that can be thought of as an extension to math. It extends the domain of the functions to allow for complex values, however we do not always use cmath even if it is more general. This is because even if we work with entirely real numbers, cmath still produced a complex output and explicity states $0j$ complex value which can complicate calculations as although mathematically there is no imaginary component python cannot distinguish this and so certain operations fail if cmath is used over math to produce output. It is possible to use cmath and take the real value with a command like $np.real()$ however it was felt this was more complicated than just using math where cmath was not explicitly required.

- **matplotlib** - This is the main package used for plotting in our analysis

- **mpmath** - Contains additional mathematical functions and abilities such as "mpmath.re()" to calculate the real component of the argument

# Appendix B

# Defining Functions in Python

Since this coding project was concerned with calculating different quantities it was important to have a consistent method of handling functions. Although Python has the downside of being poorly able to handle algebraic functions it does have the facilities for one to specify a function to produce a numerical output. To illustrate this it is perhaps easier to view this through an example. Suppose we wish to consider the equation for a straight line, $y = mx + c$, this can be calculated in python by,

```
def y(m,x,c):
a=mx
b=c
return (a+b)
```

this reads as first specifying the function we wish to define is "y" which requires m,x, and c to produce a numerical output and the combination of these used to produce an output is specified in the function definition and the output calculated inside the brackets of "return()". Of course for this simple example one could have just defined the output all inside the return function however since in general this is not practical in out analysis since the equations relevant are not usually so simple it was broken down as it would be inside the function definition where sections of the equation could be calculated separately to reduce the chance of errors when inputting the relevant expressions and to make debugging easier.

# Appendix C

# Classes in Python

In python one can define a class for a new type of object to be considered. After the class is is named it must be initialised to allow the attributes attached to the class to be activated. The method for this initialising is the same for every class however it varies in terms of what is to be initialised since not all classes will of course have the same attributes.

```
class ClassName
def __init__(self):
self.parameter1=0.
.
.
```

Above is an example of a class named "ClassName" which is initialised by setting parameter1 to a value of zero. This is the simplest case of initialisation since the only argument required for initialising is the "self" argument. In python "self" is used to represent a self contained attribute of the class. One can see that the parameter1 value is an aspect of "self" which means it is an attribute only of this class and cannot be called anywhere outside of this specific class, in other words it is not a global value. After the attributes are initialised they can be used in anyway anywhere in the class. This was mainly used in this project in terms of model classes for example,

```
class Model1process
def __init__(self):
self.mass1=0.
.
```

```
.
.
self.massn=0
def evaluate(self,list,et)
self.mass1=self.write_effective_theory(et)[0]
.
.
.
self.massn=self.write_effective_theory(et)[n]
```

is some physics model named "Model1process" which has $n$ mass parameters which are attributes of this model which are first initialised and then the value of the masses is taken from the effective theory parameter relations as outlined earlier with a general example of the class input. This approach for class definitions is used throughout for all the models considered. When applying the general code to the physics concerned here, the attributes initialised for the class are the parameters for the observables.

Combining the above gives the model class containing the input parameters,

```
def __init__(self, central):
self.para = central
self.set_parameter([])

def set_parameter(self,fitparlist):
for parlist in fitparlist:
self.para[parlist[0]] = parlist[1]
self._set_wc()
self.mass1=self.para['mass1']
.
.
.
self.massn=self.para['massn']
def write_effective_theory(self, et):
et.model = [self.mass1,...,self.massn]
return et.model
```

# Appendix D

# Compartmentalising Python Code

The python code was written using the GNU Emacs editor. This offered multiple advantages from a coding aspect, one of which was the ability to separate the code into different sections which allowed work to be done on different parts of the code independently. Since Emacs has the ability to compile in different coding languages the first step was to define a python based coding environment which is done as follows,

```
#+BEGIN_SRC python
.
.
.
.
.
#+END_SRC
```

where the relevant code is written in between BEGIN and END as indicated above. The sections can be given names to allow for easy identification without having to explicitly read through the section of code every time to determine its function. The naming is specified just before the coding environment is created,

```
#+NAME: example
#+BEGIN_SRC python
```

where “example” represents what the user would label the section. Naming has the additional advantage that it allows the section to be copied anywhere it is needed in the code without having to copy and paste it each time. In Emacs this is done by writing the name of the section in between angle brackets,

```
<<example>>
```

wherever it is needed. This has the benefit of being space saving when writing in Emacs and offers a quick and easy way to import whole chunks of code multiple times when required throughout.

As discussed the code was split into separate parts for generating points and analysing them both of which can be run independently. This was done easily in Emacs by specifying the output labelling with the name of the code section,

```
#+NAME: example
#+BEGIN_SRC python :noweb yes :tangle example.py :results output
```

which creates a python file “example.py” containing everything in the corresponding environment which is saved in the same directory as the .org file used to write the code. This also allows one to write python modules which can be imported just like any other modules as shown earlier. A python module was created like this for each model to be analysed. This is achieved by first having the model inputs outlined in section 4.1 all under one section for each model which is the exported as a python file,

```
#+NAME: examplefile
#+BEGIN_SRC python :noweb yes :tangle exampledata.py :results output
import numpy as np

<<example_model>>
#+END_SRC
```

NumPy is imported each time for the model files since each are individual separate python codes which do not contain the import section from the main code. Since the input for each model input is output to a python file in the same directory as the main .org file this allowed them to be called whenever needed (as outlined in the sections relating to calling the models for analysis and generating) by,

```
from exampledata import *
```

where the asterisk represents everything from the file "exampledata.py" being imported into the code and the code from the file which is then run. Although this largely achieves the same as using the angle brackets to copy the code sections throughout the code where needed it was decided for the models to export the inputs into a python file just so the framework allows for the possibility for the inputs to be used in different codes keeping in line with having the components of the project as general as possible.

Another advantage to this compartmentalising of the code is that it easily allows the examination of the effects of changing the model input without having to code in the observable relations each time. This methodology is indicative of the approach of the project to give the code the ability to "swap heads" meaning one can take different model inputs and use them with different calculations to create whichever combination the user requires for their analysis.

# Appendix E

# Diagonalising Matrices in Python

Diagonalising matrices was necessary for the calculations of the observables in the MSSM. This was accomplished by employing the numpy package in Python which contained the relevant tools. For some matrix $A$ one wishes to diagonalise in Python the following command is used,

a ,X=np . l i n a l g . e i g (A)

Outputs the right eigenvector and eigenvalue(s),

$$AX = Xa$$

where $X$ denotes the eigenvector and $a$ are the eigenvalues usually masses of the sparticles.

# Appendix F

# Integration in Python

The integration required for the calculations in the analysis was performed in python using the SciPy package using the command,

```
scipy.integrate.quad(function,a,b)
```

where *function* denotes the function one wishes to integrate from $b$ to $a$. Below we give an example of integration used in our calculation,

$$F_H(z) = \frac{z}{2}\int_0^1 dx \frac{1-2x(1-x)}{x(1-x)-z}\log\frac{x(1-x)}{z}, \tag{F.1}$$

$F_H(z)$ is defined as a function of $z$. To compute this integral we first specify the integrand which is set to a function of another variable, this time that which is to be integrated over, $x$. After the integrand is defined and its value returned it is input into the command as outlined above and the integral region specified. This output produces the value of the integral and the error given from the numerical integration. Since for our analysis we were only concerned with the integral value we set the output to only retain this by specifying we wanted the first element in the output only, denoted with a [0] at the end.

```
def FH(z):
    def integrand(x):
        a=(1-(2*x*(1-x)))/((x*(1-x))-z)
        b=math.log(x*(1-x)/z)
    return(a*b)
    I=scipy.integrate.quad(integrand,0,1)[0]
    return (0.5*z*I)
```

# Appendix G

# MSSM LFV Functions

The following are contributions to lepton flavour violating observables in the MSSM. These expressions are taken from [47] and are included here for completeness,

## G.1 Loop integrals

$$I_N(m_1^2,\ldots,m_N^2) \equiv \frac{i}{\pi^2}\int \frac{\mathrm{d}^4k}{(k^2-m_1^2)\ldots(k^2-m_N^2)} = (-1)^{N+1}\int_0^\infty \frac{s\,\mathrm{d}s}{(s+m_1^2)\ldots(s+m_N^2)} \tag{G.1}$$

$$J_N(m_1^2,\ldots,m_N^2) \equiv \frac{i}{\pi^2}\int \frac{k^2\,\mathrm{d}^4k}{(k^2-m_1^2)\ldots(k^2-m_N^2)} = (-1)^{N}\int_0^\infty \frac{s^2\,\mathrm{d}s}{(s+m_1^2)\ldots(s+m_N^2)} \tag{G.2}$$

$$K_N(m_1^2,\ldots,m_N^2) \equiv \frac{i}{\pi^2}\int \frac{k^4\,\mathrm{d}^4k}{(k^2-m_1^2)\ldots(k^2-m_N^2)} = (-1)^{N+1}\int_0^\infty \frac{s^3\,\mathrm{d}s}{(s+m_1^2)\ldots(s+m_N^2)}. \tag{G.3}$$

## G.2 $D^\gamma_{L,R}$ Contribution

$D^\gamma_{L,R}$ is decomposed as $D^\gamma_{L,R} = D^{\gamma(a)}_{L,R} + D^{\gamma(b)}_{L,R} + D^{\gamma(c)}_{L,R}$.

$$\frac{D_L^{\gamma(a)}}{s_L c_L} = \frac{g^2}{16\pi^2}\cdot\frac{1}{4}\left[\tilde{m}^2_{L_2} J_5(M_2^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2}) - 2M_2^2 J_5(M_2^2,M_2^2,M_2^2,M_2^2,\tilde{m}^2_{L_2})\right]$$
$$+\frac{g'^2}{16\pi^2}\cdot\frac{1}{4}\tilde{m}^2_{L_2} J_5(M_1^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2}) - (L_2\to L_3) \tag{G.4}$$
$$\frac{D_R^{\gamma(a)}}{s_R c_R} = \frac{g'^2}{16\pi^2}\tilde{m}^2_{R_2} J_5(M_1^2,\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{R_2}) - (R_2\to R_3) \tag{G.5}$$

$$\frac{D_L^{\gamma(b)}}{s_L c_L} = \frac{g^2}{16\pi^2}\cdot\frac{1}{2}\left[-\tilde{m}^2_{L_2} J_5(M_2^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2}) - 2\tilde{m}^4_{L_2} I_5(M_2^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2})\right.$$
$$+M_2\mu\tan\beta\left(\tilde{m}^2_{L_2} I_5(M_2^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2}) + 2J_5(M_2^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2})\right.$$
$$\left.\left.+2J_5(M_2^2,M_2^2,M_2^2,\mu^2,\tilde{m}^2_{L_2}) + 2J_5(M_2^2,\mu^2,\mu^2,\mu^2,\tilde{m}^2_{L_2})\right)\right]$$
$$+\frac{g'^2}{16\pi^2}\frac{\tilde{m}^2_{L_2}}{2}\left[J_5(M_1^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2}) - M_1\mu\tan\beta\, I_5(M_1^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{L_2})\right]$$
$$-(L_2\to L_3) \tag{G.6}$$
$$\frac{D_R^{\gamma(b)}}{s_R c_R} = \frac{g'^2}{16\pi^2}\tilde{m}^2_{R_2}\left[-J_5(M_1^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{R_2}) + M_1\mu\tan\beta\, I_5(M_1^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{R_2})\right]$$
$$-(R_2\to R_3) \tag{G.7}$$

$$D_L^{\gamma(c)} = -\frac{g'^2}{16\pi^2}M_1^3\times$$
$$\left\{\left[s_L c_L\left(s_R^2[A_\tau+\mu\tan\beta]+s_R c_R A^R_{\mu\tau}\right)+c_L^2 s_R^2 A^L_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_2},\tilde{m}^2_{R_2})\right.$$
$$-\left[s_L c_L\left(s_R^2[A_\tau+\mu\tan\beta]+s_R c_R A^R_{\mu\tau}\right)-s_L^2 s_R^2 A^L_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_3},\tilde{m}^2_{R_2})$$
$$+\left[s_L c_L\left(c_R^2[A_\tau+\mu\tan\beta]-s_R c_R A^R_{\mu\tau}\right)+c_L^2 c_R^2 A^L_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_2},\tilde{m}^2_{R_3})$$
$$\left.-\left[s_L c_L\left(c_R^2[A_\tau+\mu\tan\beta]-s_R c_R A^R_{\mu\tau}\right)-s_L^2 c_R^2 A^L_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_3},\tilde{m}^2_{R_3})\right\} \tag{G.8}$$

$$D_R^{\gamma(c)} = -\frac{g'^2}{16\pi^2}M_1^3\times$$
$$\left\{\left[s_R c_R\left(s_L^2[A_\tau+\mu\tan\beta]+s_L c_L A^L_{\mu\tau}\right)+c_R^2 s_L^2 A^R_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_2},\tilde{m}^2_{R_2})\right.$$
$$-\left[s_R c_R\left(s_L^2[A_\tau+\mu\tan\beta]+s_L c_L A^L_{\mu\tau}\right)-s_R^2 s_L^2 A^R_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_2},\tilde{m}^2_{R_3})$$
$$+\left[s_R c_R\left(c_L^2[A_\tau+\mu\tan\beta]-s_L c_L A^L_{\mu\tau}\right)+c_R^2 c_L^2 A^R_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_3},\tilde{m}^2_{R_2})$$
$$\left.-\left[s_R c_R\left(c_L^2[A_\tau+\mu\tan\beta]-s_L c_L A^L_{\mu\tau}\right)-s_R^2 c_L^2 A^R_{\mu\tau}\right] I_5(M_1^2,M_1^2,M_1^2,\tilde{m}^2_{L_3},\tilde{m}^2_{R_3})\right\} \tag{G.9}$$

## G.3 $A^Z_{L,R}$ Contribution

$A^Z_{L,R}$ is decomposed as $D^Z_{L,R} = A^{Z(a)}_{L,R} + A^{Z(b)}_{L,R} + A^{Z(c)}_{L,R}$.

$$
\begin{aligned}
\frac{A_L^{Z(a)}}{s_L c_L} &= \frac{g^2 c_W^2}{16\pi^2}\cdot\frac{1}{8}\left[-(2+3c_{2\beta})\left(\mu^2 J_5(M_2^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) + 2J_4(M_2^2,M_2^2,\mu^2,\tilde{m}^2_{L_2})\right)\right.\\
&-(2-3c_{2\beta})M_2^2\left(\mu^2 I_5(M_2^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) - I_4(M_2^2,M_2^2,\mu^2,\tilde{m}^2_{L_2})\right)\\
&\left.+4s_{2\beta}\mu M_2\, J_5(M_2^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2})\right]\\
&+\frac{g'^2 c_W^2}{16\pi^2}\cdot\frac{1}{4}\,c_{2\beta}\left[\mu^2 J_5(M_1^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) + 2J_4(M_1^2,M_2^2,\mu^2,\tilde{m}^2_{L_2})\right.\\
&\left.-M_1M_2\left(\mu^2 I_5(M_1^2,M_2^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) - I_4(M_1^2,M_2^2,\mu^2,\tilde{m}^2_{L_2})\right)\right]\\
&+\frac{g'^2 s_W^2}{16\pi^2}\cdot\frac{1}{8}\,c_{2\beta}\left[-\mu^2 J_5(M_1^2,M_1^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) - 2J_4(M_1^2,M_1^2,\mu^2,\tilde{m}^2_{L_2})\right.\\
&\left.+M_1^2\left(\mu^2 I_5(M_1^2,M_1^2,\mu^2,\mu^2,\tilde{m}^2_{L_2}) - I_4(M_1^2,M_1^2,\mu^2,\tilde{m}^2_{L_2})\right)\right]\\
&-(L_2\to L_3) \qquad\qquad (G.10)\\
\frac{A_R^{Z(a)}}{s_R c_R} &= \frac{g'^2 s_W^2}{16\pi^2}\cdot\frac{1}{2}\,c_{2\beta}\left[\mu^2 J_5(M_1^2,M_1^2,\mu^2,\mu^2,\tilde{m}^2_{R_2}) + 2J_4(M_1^2,M_1^2,\mu^2,\tilde{m}^2_{R_2})\right.\\
&\left.-M_1^2\left(\mu^2 I_5(M_1^2,M_1^2,\mu^2,\mu^2,\tilde{m}^2_{R_2}) - I_4(M_1^2,M_1^2,\mu^2,\tilde{m}^2_{R_2})\right)\right]\\
&-(R_2\to R_3) \qquad\qquad (G.11)
\end{aligned}
$$

$$
\begin{aligned}
\frac{A_L^{Z(b)}}{s_L c_L} &= \frac{h_\tau^2 c_W^2}{16\pi^2}\cdot\frac{1}{4}\,s_\beta^2\mu^2\left[s_R^2\left(J_5(M_2^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{L_2}) + J_5(M_2^2,\mu^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{L_2})\right)\right.\\
&\left.+c_R^2\left(J_5(M_2^2,\mu^2,\tilde{m}^2_{R_3},\tilde{m}^2_{R_3},\tilde{m}^2_{L_2}) + J_5(M_2^2,\mu^2,\mu^2,\tilde{m}^2_{R_3},\tilde{m}^2_{L_2})\right)\right]\\
&+\frac{h_\tau^2 s_W^2}{16\pi^2}\cdot\frac{1}{4}\,s_\beta^2\mu^2\left[-s_R^2\left(J_5(M_1^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{R_2},\tilde{m}^2_{L_2}) + J_5(M_1^2,\mu^2,\mu^2,\tilde{m}^2_{R_2},\tilde{m}^2_{L_2})\right)\right.\\
&\left.-c_R^2\left(J_5(M_1^2,\mu^2,\tilde{m}^2_{R_3},\tilde{m}^2_{R_3},\tilde{m}^2_{L_2}) + J_5(M_1^2,\mu^2,\mu^2,\tilde{m}^2_{R_3},\tilde{m}^2_{L_2})\right)\right]\\
&-(L_2\to L_3) \qquad\qquad (G.12)\\
\frac{A_R^{Z(b)}}{s_R c_R} &= \frac{h_\tau^2 s_W^2}{16\pi^2}\cdot\frac{1}{2}\,s_\beta^2\mu^2\left[-s_L^2\left(J_5(M_1^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{L_2},\tilde{m}^2_{R_2}) + J_5(M_1^2,\mu^2,\mu^2,\tilde{m}^2_{L_2},\tilde{m}^2_{R_2})\right)\right.\\
&\left.-c_L^2\left(J_5(M_1^2,\mu^2,\tilde{m}^2_{L_3},\tilde{m}^2_{L_3},\tilde{m}^2_{R_2}) + J_5(M_1^2,\mu^2,\mu^2,\tilde{m}^2_{L_3},\tilde{m}^2_{R_2})\right)\right]\\
&-(R_2\to R_3) \qquad\qquad (G.13)
\end{aligned}
$$

$$
\begin{aligned}
\frac{A_L^{Z(c)}}{s_L c_L} = & \frac{h_\tau^2 c_W^2}{16\pi^2}\cdot\frac{1}{4}\, s_\beta^2\mu^2\left[s_L^2\left(s_R^2 J_5(M_2^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right.\right.\\
& +c_R^2 J_5(M_2^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\Big) - c_L^2\left(s_R^2 J_5(M_2^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right.\\
& \left.+c_R^2 J_5(M_2^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right) - (s_L^2-c_L^2)\left(s_R^2 J_5(M_2^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right.\\
& \left.+c_R^2 J_5(M_2^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right)\Big]\\
& + \frac{h_\tau^2 s_W^2}{16\pi^2}\cdot\frac{1}{4}\, s_\beta^2\mu^2\left[s_L^2\left(s_R^2 J_5(M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right.\right.\\
& +c_R^2 J_5(M_1^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\Big) - c_L^2\left(s_R^2 J_5(M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right.\\
& \left.+c_R^2 J_5(M_1^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right) - (s_L^2-c_L^2)\left(s_R^2 J_5(M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right.\\
& \left.+c_R^2 J_5(M_1^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right)\Big] \qquad \text{(G.14)}\\
\frac{A_R^{Z(c)}}{s_R c_R} = & \frac{h_\tau^2 s_W^2}{16\pi^2} s_\beta^2\mu^2\left[-s_R^2\left(s_L^2 J_5(M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2)\right.\right.\\
& +c_L^2 J_5(M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2)\Big) + c_R^2\left(s_L^2 J_5(M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2)\right.\\
& \left.+c_L^2 J_5(M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2)\right) + (s_R^2-c_R^2)\left(s_L^2 J_5(M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_3}^2)\right.\\
& \left.+c_L^2 J_5(M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_3}^2)\right)\Big] \qquad \text{(G.15)}
\end{aligned}
$$

## G.4 Contributions to $C_{L,R}^{\gamma}$

$$
\begin{aligned}
\frac{C_L^\gamma}{s_L c_L} = & \frac{g^2}{16\pi^2}\cdot\frac{1}{12}\left[-K_5(M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right.\\
& \left.-4K_5(M_2^2,M_2^2,M_2^2,M_2^2,\tilde{m}_{L_2}^2)+6M_2^2 J_5(M_2^2,M_2^2,M_2^2,M_2^2,\tilde{m}_{L_2}^2)\right]\\
& + \frac{g'^2}{16\pi^2}\cdot\frac{1}{12}[-K_5(M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)] - (L_2\to L_3) \qquad \text{(G.16)}\\
\frac{C_R^\gamma}{s_R c_R} = & \frac{g'^2}{16\pi^2}\cdot\frac{1}{3}\left[-K_5(M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2)\right] - (R_2\to R_3) \qquad \text{(G.17)}
\end{aligned}
$$

## G.5 $B_{L,R}^{f_{L,R}}$ Contribution

For $f \neq \mu, \tau$:

$$\begin{aligned}
\frac{B_L^{f_L}}{s_L c_L} &= \frac{g^4}{16\pi^2} \cdot \frac{1}{16} \Big[ - \left(1 + 4\,\delta_{T^3_{f_L},-\frac{1}{2}}\right) J_4(M_2^2, M_2^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) \\
&\quad -2 \left(1 + 4\,\delta_{T^3_{f_L},\frac{1}{2}}\right) M_2^2 I_4(M_2^2, M_2^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) \Big] \\
&\quad + \frac{g^2 g'^2}{16\pi^2} \cdot \frac{1}{2} (-T^3_{f_L} Y_{f_L}) \left[ J_4(M_1^2, M_2^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) + 2 M_1 M_2 I_4(M_1^2, M_2^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) \right] \\
&\quad + \frac{g'^4}{16\pi^2} \cdot \frac{1}{4} (-Y^2_{f_L}) \left[ J_4(M_1^2, M_1^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) + 2 M_1^2 I_4(M_1^2, M_1^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_L}) \right] \\
&\quad - (L_2 \to L_3) \qquad \text{(G.18)} \\
\frac{B_L^{f_R}}{s_L c_L} &= \frac{g'^4}{16\pi^2} \cdot \frac{1}{4} Y^2_{f_R} \left[ J_4(M_1^2, M_1^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_R}) + 2 M_1^2 I_4(M_1^2, M_1^2, \tilde{m}^2_{L_2}, \tilde{m}^2_{\tilde{f}_R}) \right] \\
&\quad - (L_2 \to L_3) \qquad \text{(G.19)} \\
\frac{B_R^{f_L}}{s_R c_R} &= \frac{g'^4}{16\pi^2} Y^2_{f_L} \left[ J_4(M_1^2, M_1^2, \tilde{m}^2_{R_2}, \tilde{m}^2_{\tilde{f}_L}) + 2 M_1^2 I_4(M_1^2, M_1^2, \tilde{m}^2_{R_2}, \tilde{m}^2_{\tilde{f}_L}) \right] \\
&\quad - (R_2 \to R_3) \qquad \text{(G.20)} \\
\frac{B_R^{f_R}}{s_R c_R} &= \frac{g'^4}{16\pi^2} (-Y^2_{f_R}) \left[ J_4(M_1^2, M_1^2, \tilde{m}^2_{R_2}, \tilde{m}^2_{\tilde{f}_R}) + 2 M_1^2 I_4(M_1^2, M_1^2, \tilde{m}^2_{R_2}, \tilde{m}^2_{\tilde{f}_R}) \right] \\
&\quad - (R_2 \to R_3) \qquad \text{(G.21)}
\end{aligned}$$

For $f = \mu$:

$$\begin{aligned}
\frac{B_L^{\mu_L}}{s_L c_L} = \frac{g^4}{16\pi^2}\cdot\frac{1}{16}\big[&-c_L^2\left(5J_4(M_2^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)+2M_2^2I_4(M_2^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right)\\
&+s_L^2\left(5J_4(M_2^2,M_2^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)+2M_2^2I_4(M_2^2,M_2^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right)\\
&+(c_L^2-s_L^2)\left(5J_4(M_2^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)+2M_2^2I_4(M_2^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right)\big]\\
+\frac{g^2g'^2}{16\pi^2}\cdot\frac{1}{8}\big[&-c_L^2\left(J_4(M_1^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)+2M_1M_2I_4(M_1^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right)\\
&+s_L^2\left(J_4(M_1^2,M_2^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)+2M_1M_2I_4(M_1^2,M_2^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right)\\
&+(c_L^2-s_L^2)\left(J_4(M_1^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)+2M_1M_2I_4(M_1^2,M_2^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right)\big]\\
+\frac{g'^4}{16\pi^2}\cdot\frac{1}{16}\big[&-c_L^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_2}^2)\right)\\
&+s_L^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{L_3}^2)\right)\\
&+(c_L^2-s_L^2)\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{L_3}^2)\right)\big]
\end{aligned} \tag{G.22}$$

$$\begin{aligned}
\frac{B_L^{\mu_R}}{s_L c_L} = \frac{g'^4}{16\pi^2}\cdot\frac{1}{4}\big[&c_R^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2)\right)\\
&+s_R^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_3}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_3}^2)\right)\big]\\
&-(L_2\to L_3)
\end{aligned} \tag{G.23}$$

$$\begin{aligned}
\frac{B_R^{\mu_L}}{s_R c_R} = \frac{g'^4}{16\pi^2}\cdot\frac{1}{4}\big[&c_L^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_2}^2,\tilde{m}_{R_2}^2)\right)\\
&+s_L^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{R_2}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{L_3}^2,\tilde{m}_{R_2}^2)\right)\big]\\
&-(R_2\to R_3)
\end{aligned} \tag{G.24}$$

$$\begin{aligned}
\frac{B_R^{\mu_R}}{s_R c_R} = \frac{g'^4}{16\pi^2}\big[&-c_R^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_2}^2)\right)\\
&+s_R^2\left(J_4(M_1^2,M_1^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{R_3}^2,\tilde{m}_{R_3}^2)\right)\\
&+(c_R^2-s_R^2)\left(J_4(M_1^2,M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_3}^2)+2M_1^2I_4(M_1^2,M_1^2,\tilde{m}_{R_2}^2,\tilde{m}_{R_3}^2)\right)\big]
\end{aligned} \tag{G.25}$$

# Appendix H

# Function Definitions for $a_\mu^{non-Yuk}$ in the 2HDM

The $\Phi$ loop function is defined in Ref. [84] and we give this definition below,

$$\Phi(m_1,m_2,m_3) = \frac{\lambda}{2}\Bigg[2\ln(\alpha_+)\ln(\alpha_-) - \ln\left(\frac{m_1^2}{m_3^2}\right)\ln\left(\frac{m_2^2}{m_3^2}\right) - 2\mathrm{Li}_2(\alpha_+) - 2\mathrm{Li}_2(\alpha_-) + \frac{\pi^2}{3}\Bigg], \tag{H.1}$$

$$\lambda = \sqrt{m_1^4 + m_2^4 + m_3^4 - 2m_1^2m_2^2 - 2m_2^2m_3^2 - 2m_3^2m_1^2}\,, \tag{H.2}$$

$$\alpha_\pm = \frac{m_3^2 \pm m_1^2 \mp m_2^2 - \lambda}{2m_3^2}. \tag{H.3}$$

The abbreviations appearing in $a_\mu^{\text{non-Yuk}}$ are [51].

$$\mathcal{T}_0(u,\omega) = \frac{9}{c_{\mathrm{W}}^4}\frac{(u-\omega)(c_{\mathrm{W}}^2(u-\omega)(u+2\omega) - (u-\omega)^3 + c_{\mathrm{W}}^4\omega)}{c_{\mathrm{W}}^4 + (u-\omega)^2 - 2c_{\mathrm{W}}^2(u+\omega)} \times \Phi(\sqrt{u},\sqrt{\omega},c_{\mathrm{W}}), \tag{H.4}$$

$$\mathcal{T}_1(u,\omega) = \frac{9}{c_{\mathrm{W}}^4}(u-\omega)(c_{\mathrm{W}}^2\omega - (u-\omega)^2)\mathrm{Li}_2(1-u/\omega), \tag{H.5}$$

$$\mathcal{T}_2^\pm(u,\omega) = \ln(u)\left(\frac{6u^2 + c_{\mathrm{W}}^2(u-x_{H^\pm}) + 2c_{\mathrm{W}}^4(u-x_{H^\pm})}{2(u-\omega)} + f_6\frac{(u-x_{H^\pm})^2(3c_{\mathrm{W}}^4 + 3c_{\mathrm{W}}^2(u-x_{H^\pm}) + (u-x_{H^\pm})^2}{c_{\mathrm{W}}^2(u-\omega)} \pm f_7\frac{3u^2(u-x_{H^\pm})}{(x_A-x_H)(u-\omega)} - f_8\frac{3u(u-x_{H^\pm})^2}{2(u-\omega)}\right.$$

$$- f_9\frac{3u(u-x_{H^\pm})}{2(u-\omega)}\Bigg), \tag{H.6}$$

$$\mathcal{T}_4(u,\omega) = \frac{(u-\omega)\ln(u)}{4} f_5(x_A(3+2x_H) - x_A^2 + 3x_H - x_H^2 - 3), \tag{H.7}$$

$$\begin{aligned}\mathcal{T}_5(u,\omega) =& \ln(u)\left(\frac{3}{2}u + \frac{f_6}{c_{\rm W}^2}((u-\omega)^3 + 3c_{\rm W}^2(u-\omega)^2 + 3c_{\rm W}^4(u-\omega))\right.\\ &\left. -\frac{3}{2}f_8 u(u-\omega) - \frac{c_{\rm W}^2}{2} - c_{\rm W}^4\right),\end{aligned} \tag{H.8}$$

$$\begin{aligned}\mathcal{T}_6(u,\omega) =& \frac{9}{2}\left(\frac{(u-\omega)(u^2-2u\omega+\omega(\omega-c_{\rm W}^2))}{c_{\rm W}^4}\ln\left(\frac{u}{\omega}\right)\ln\left(\frac{\omega}{c_{\rm W}^2}\right)\right.\\ &\left. +\frac{\ln(c_{\rm W}^2)}{c_{\rm W}^2}(2u^2 + u(c_{\rm W}^2 - 4\omega) - \omega(c_{\rm W}^2 - 2\omega))\right),\end{aligned} \tag{H.9}$$

$$\begin{aligned}\mathcal{T}_7(u,\omega) =& -0.5 f_5(2(u+\omega) - (u-\omega)^2 - 1)\ln\left(\frac{\mathcal{S}_1(u,\omega)}{2\sqrt{u\omega}}\right)\\ &\times\left(u+\omega-1-\frac{4u\omega}{\mathcal{S}_1(u,\omega)}\right),\end{aligned} \tag{H.10}$$

$$\mathcal{S}_1(u,\omega) = u+\omega-1+\sqrt{1+(u-\omega)^2-2(u+\omega)}, \tag{H.11}$$

$$\begin{aligned}\mathcal{T}_8(u,\omega) =& 2f_6(4u\omega - (u+\omega-c_{\rm W}^2)^2)\ln\left(\frac{\mathcal{S}_2(u,\omega)}{2\sqrt{u\omega}}\right)\\ &\times\left(\frac{(u+\omega)}{c_{\rm W}^2} - \frac{4u\omega}{c_{\rm W}^2\mathcal{S}_2(u,\omega)} - 1\right),\end{aligned} \tag{H.12}$$

$$\mathcal{S}_2(u,\omega) = u+\omega-c_{\rm W}^2+\sqrt{(u+\omega-c_{\rm W}^2)^2-4u\omega}. \tag{H.13}$$

Note for our analysis equation H.10 differs by a factor of $-\frac{1}{2}$ from Ref.[51] to allow our analysis to produce consistent output with literature when tested the corresponding inputs.

$$f_1 = \frac{7}{2} - \frac{25}{2c_{\rm W}^2} + 4c_{\rm W}^2 - 4c_{\rm W}^4 = -12, \tag{H.14}$$

$$f_2 = 2(17 - 24c_{\rm W}^2 + 56c_{\rm W}^4 - 128c_{\rm W}^6 + 64c_{\rm W}^8) = -9.1, \tag{H.15}$$

$$f_3 = \frac{25 - 32c_{\rm W}^2 + 4c_{\rm W}^4}{c_{\rm W}^2 s_{\rm W}^2} = 15, \tag{H.16}$$

$$f_4 = \frac{13}{2} - 15c_{\rm W}^2 + 10c_{\rm W}^4 = -0.9, \tag{H.17}$$

$$f_5 = \frac{c_{\rm W}^2(5 - 16c_{\rm W}^2 + 8c_{\rm W}^4)}{s_{\rm W}^2} = -9, \tag{H.18}$$

$$f_6 = \frac{7 - 14c_{\rm W}^2 + 4c_{\rm W}^4}{4c_{\rm W}^2 s_{\rm W}^2} = -2, \tag{H.19}$$

$$f_7 = 1 - 6c_{\mathrm{W}}^2 + 4c_{\mathrm{W}}^4 = -1.2, \tag{H.20}$$

$$f_8 = \frac{13 - 20c_{\mathrm{W}}^2 + 4c_{\mathrm{W}}^4}{c_{\mathrm{W}}^2 s_{\mathrm{W}}^2} = -0.7, \tag{H.21}$$

$$f_9 = 7 - 12c_{\mathrm{W}}^2 + 8c_{\mathrm{W}}^4 = 2.5. \tag{H.22}$$

# Appendix I

# Functions used in the Calculation of the Yukawa dependent component of $a_\mu$ in the 2HDM

Below we list the list the expressions of the coefficients used to calculate the Yukawa dependent of $a_\mu$ given in equation 7.23. These expressions are taken from Ref. [51],

$$a^0_{0,0} = b(x_{H_{\rm SM}}, x_{H^\pm})\mathcal{F}^0_m(x_{H_{\rm SM}}, x_{H^\pm}); \tag{I.1}$$

$$a^0_{0,z} = -b(x_H, 0)\left[\mathcal{F}^0_m(x_H, x_{H^\pm}) + \mathcal{F}^\pm_m(x_H, x_{H^\pm})\right]; \tag{I.2}$$

$$a^0_{5,0} = \mathcal{F}^0_m(x_{H_{\rm SM}}, x_{H^\pm}); \tag{I.3}$$

$$a^0_{5,z} = -\frac{1}{2}\left[\mathcal{F}^0_m(x_H, x_{H^\pm}) + \mathcal{F}^\pm_m(x_H, x_{H^\pm})\right]; \tag{I.4}$$

$$a^1_{0,0} = b(x_H, 0)\mathcal{F}^0_m(x_H, x_{H^\pm}) - (x_H \to x_{H_{\rm SM}}); \tag{I.5}$$

$$a^1_{0,z} = -\Big[b(x_H, x_{H^\pm})\left(\mathcal{F}^0_m(x_H, x_{H^\pm}) + \mathcal{F}^\pm_m(x_H, x_{H^\pm})\right) - \mathcal{F}_3(x_H, x_{H^\pm}) - (x_H \to x_{H_{\rm SM}})\Big] + \mathcal{F}_2(x_H); \tag{I.6}$$

$$a^1_{5,0} = \frac{\mathcal{F}^0_m(x_H, x_{H^\pm})}{2} - (x_H \to x_{H_{\rm SM}}); \tag{I.7}$$

$$a^1_{5,z} = -\mathcal{F}^0_m(x_H, x_{H^\pm}) - \mathcal{F}^\pm_m(x_H, x_{H^\pm}) - (x_H \to x_{H_{\rm SM}}). \tag{I.8}$$

The constituents of the above expressions of the coefficients are given below again taken from Ref. [51]

$$b(u,\omega) = \frac{\alpha\pi}{c_{\rm W}^2(-1+c_{\rm W}^2)}(u+2\omega), \tag{I.9}$$

$$\mathcal{F}_m^0(u,\omega) = \frac{\alpha_{em}^2}{576\,\pi^2\,c_{\rm W}^4 s_{\rm W}^4}\,\frac{m_\mu^2}{M_Z^2}\,\left(\frac{1}{\alpha\pi}\frac{c_{\rm W}^2(-1+c_{\rm W}^2)}{(u+2\omega)}\right)\mathcal{F}_1(u,\omega), \tag{I.10}$$

$$\mathcal{F}_m^{\pm}(u,\omega) = \frac{\alpha_{em}^2}{576\,\pi^2\,c_{\rm W}^4 s_{\rm W}^4}\,\frac{m_\mu^2}{M_Z^2}\,\left(-\frac{9(-1+c_{\rm W}^2)}{\alpha\pi}\right)\left(\frac{\mathcal{T}_9(u,\omega)}{2}+\mathcal{T}_{10}(u,\omega)\right), \tag{I.11}$$

$$\begin{aligned}\mathcal{F}_1(u,\omega) =& -72c_{\rm W}^2(-1+c_{\rm W}^2)\frac{u+2\omega}{u} - 36c_{\rm W}^2(-1+c_{\rm W}^2)\frac{u+2\omega}{u}\ln(\omega)\\ &+9(-8c_{\rm W}^4-3u+2c_{\rm W}^2(4+u))\frac{(u+2\omega)}{2(u-1)u}\ln(u)\\ &-9(3-10c_{\rm W}^2+8c_{\rm W}^4)\frac{\omega(u+2\omega)}{(4\omega-1)(u-1)}\Phi(\sqrt{\omega},\sqrt{\omega},1)\\ &+9(8c_{\rm W}^4+3u-2c_{\rm W}^2(4+u))\frac{\omega(u+2\omega)}{(4\omega-u)(u-1)u^2}\Phi(\sqrt{u},\sqrt{\omega},\sqrt{\omega}),\end{aligned} \tag{I.12}$$

$$\begin{aligned}\mathcal{T}_9(u,\omega) =& -\frac{2\left(c_{\rm W}^4\omega+c_{\rm W}^2\left(u^2+u\omega-2\omega^2\right)-(u-\omega)^3\right)\Phi\left(\sqrt{u},\sqrt{\omega},c_{\rm W}\right)}{\left(c_{\rm W}^2-\omega\right)\left(c_{\rm W}^4-2c_{\rm W}^2(u+\omega)+(u-\omega)^2\right)}\\ &+\frac{2c_{\rm W}^4\left(u^2-4u\omega+2\omega^2\right)\Phi\left(\sqrt{u},\sqrt{\omega},\sqrt{\omega}\right)}{\omega^2\left(\omega-c_{\rm W}^2\right)(u-4\omega)}\\ &-\frac{2\left(c_{\rm W}^2u(u-2\omega)+\omega(u-\omega)^2\right)\mathrm{Li}_2\left(1-\frac{u}{\omega}\right)}{\omega^2},\end{aligned} \tag{I.13}$$

$$\begin{aligned}\mathcal{T}_{10}(u,\omega) =& \frac{u^2-c_{\rm W}^2\omega-2u\omega+\omega^2}{2(c_{\rm W}^2-\omega)}\ln\left(\frac{\omega}{u}\right)\ln\left(\frac{\omega}{c_{\rm W}^2}\right)\\ &+\frac{c_{\rm W}^2(c_{\rm W}^2+2u-2\omega)}{2(c_{\rm W}^2-\omega)}\ln\left(\frac{\omega}{c_{\rm W}^2}\right)+\frac{c_{\rm W}^2}{\omega}u\ln\left(\frac{\omega}{u}\right)+\frac{c_{\rm W}^2}{\omega}(\omega-u),\end{aligned} \tag{I.14}$$

$$\begin{aligned}\mathcal{F}_2(u) =& F^{\rm W}(u)+F^{\rm Z}(u)+\\ &+\frac{\alpha_{em}^2}{576\,\pi^2\,c_{\rm W}^4 s_{\rm W}^4}\,\frac{m_\mu^2}{M_Z^2}\left\{\frac{8c_{\rm W}^6\pi^2}{u^2}+\frac{F_0}{u}+\frac{393c_{\rm W}^2}{8}\right.\\ &+\left(\frac{F_1}{u}+F_2+F_3u\right)\frac{\ln(c_{\rm W}^2)}{(4c_{\rm W}^2-1)(4c_{\rm W}^2-u)}\\ &+\left(\frac{F_4}{u}+F_5+F_6u+F_7u^2\right)\frac{\ln(u)}{(u-1)(4c_{\rm W}^2-u)}\end{aligned}$$

$$
\begin{aligned}
&- \frac{3}{2}\left(\frac{32c_{\rm W}^6}{u^2}+\frac{21c_{\rm W}^4}{u}+15c_{\rm W}^2-35u\right)\mathrm{Li}_2\left(1-\frac{u}{c_{\rm W}^2}\right)\\
&+(F_8+F_9u)\frac{9c_{\rm W}^2(-3+4c_{\rm W}^2)}{2}\frac{\Phi\left(c_{\rm W},c_{\rm W},1\right)}{(4c_{\rm W}^2-1)^2(u-1)}\\
&+\left[\frac{F_{10}}{u^2}+\frac{F_{11}}{u}+F_{12}+F_{13}u+F_{14}u^2+\frac{105u^3}{2}\right]\frac{\Phi\left(\sqrt{u},c_{\rm W},c_{\rm W}\right)}{(4c_{\rm W}^2-u)^2(u-1)}\Bigg\},
\end{aligned}
\tag{I.15}
$$

Below are values taken from Ref. [51] which are functions of the Weinberg angle. Their approximate values are also given assuming the standard value for the Weinberg angle, however since our analysis was focussing on examining different models and the impact of changes to the frameworks we elected to implement these values as functions of the Weinberg angle rather than just code in the values given on the right to allow analysis to easily be performed if a framework was considered with a different definition of the Weinberg angle.

$$
\begin{aligned}
F_0 =& \frac{3c_{\rm W}^4\left(-640+576c_{\rm W}^2+7\pi^2\right)}{4} && = -55.9, && \text{(I.16)}\\
F_1 =& 96c_{\rm W}^6\left(11-53c_{\rm W}^2+36c_{\rm W}^4\right) && = -380, && \text{(I.17)}\\
F_2 =& -\frac{3}{4}c_{\rm W}^2\left(-66c_{\rm W}^2-48c_{\rm W}^4+672c_{\rm W}^6\right) && = -137, && \text{(I.18)}\\
F_3 =& -\frac{3}{4}c_{\rm W}^2\left(109-430c_{\rm W}^2+120c_{\rm W}^4\right) && = +88.8, && \text{(I.19)}\\
F_4 =& 96c_{\rm W}^6\left(-11+9c_{\rm W}^2\right) && = -180, && \text{(I.20)}\\
F_5 =& \frac{45c_{\rm W}^4}{2}+192c_{\rm W}^6 && = +103, && \text{(I.21)}\\
F_6 =& \frac{3}{4}c_{\rm W}^2\left(157+90c_{\rm W}^2\right) && = +132, && \text{(I.22)}\\
F_7 =& -\frac{3}{4}\left(18+61c_{\rm W}^2\right) && = -49.0, && \text{(I.23)}\\
F_8 =& \left(-7+61c_{\rm W}^2-162c_{\rm W}^4+96c_{\rm W}^6\right) && = -12.3, && \text{(I.24)}\\
F_9 =& (1-5c_{\rm W}^2+10c_{\rm W}^4) && = +3.15, && \text{(I.25)}\\
F_{10} =& -1728c_{\rm W}^8\left(-1+c_{\rm W}^2\right) && = +140, && \text{(I.26)}\\
F_{11} =& 3c_{\rm W}^6\left(-899+768c_{\rm W}^2\right) && = -425, && \text{(I.27)}\\
F_{12} =& \left(387c_{\rm W}^4-363c_{\rm W}^6\right) && = +63.4, && \text{(I.28)}\\
F_{13} =& \frac{9}{2}c_{\rm W}^2\left(57+106c_{\rm W}^2\right) && = +486, && \text{(I.29)}\\
F_{14} =& -\frac{15}{2}\left(7+45c_{\rm W}^2\right) && = -314, && \text{(I.30)}
\end{aligned}
$$

$$
\begin{aligned}
F^{\mathrm{Z}}(u) =& \frac{\alpha_{em}^2}{576\,\pi^2\,c_{\mathrm{W}}^4 s_{\mathrm{W}}^4}\,\frac{m_\mu^2}{M_Z^2}\,\Bigg\{ Z_1 u \mathrm{Li}_2\left(1-u\right) \\
&+ \frac{Z_2}{2u^2}\Big[6(-4+u)u + \pi^2(4+3u) + 6u(4+u)\ln(u) \\
&- 6(4+3u)\mathrm{Li}_2\left(1-u\right) + 6u(2+u)\Phi\left(\sqrt{u},1,1\right)\Big] \\
&+ Z_3 u\Big[6 + \pi^2(-4+u)u + 3\ln(u)(4+(-4+u)u\ln(u)) \\
&+ 12(-4+u)u\mathrm{Li}_2\left(1-u\right) + 6(-2+u)\Phi\left(\sqrt{u},1,1\right)\Big], \qquad \text{(I.31)}
\end{aligned}
$$

$$
\begin{aligned}
Z_1 =& 3(17 - 48c_{\mathrm{W}}^2 + 32c_{\mathrm{W}}^4) && = -2.9, && \text{(I.32)}\\
Z_2 =& \left(5 - 12c_{\mathrm{W}}^2 + 8c_{\mathrm{W}}^4\right) && = 0.50, && \text{(I.33)}\\
Z_3 =& 3\left(1 - 3c_{\mathrm{W}}^2 + 2c_{\mathrm{W}}^4\right) && = -0.37, && \text{(I.34)}
\end{aligned}
$$

$$
\begin{aligned}
F^{\mathrm{W}}(u) =& \frac{\alpha_{em}^2}{576\,\pi^2\,c_{\mathrm{W}}^4 s_{\mathrm{W}}^4}\,\frac{m_\mu^2}{M_Z^2}\,\Bigg\{ -\frac{57c_{\mathrm{W}}^2}{2} - \frac{4c_{\mathrm{W}}^6\pi^2}{u^2} + \frac{3c_{\mathrm{W}}^4\left(32-3\pi^2\right)}{4u} \\
&+ \frac{3\left(16c_{\mathrm{W}}^6 + 9c_{\mathrm{W}}^4 u + 12c_{\mathrm{W}}^2 u^2 - 19u^3\right)\mathrm{Li}_2\left(1-\frac{u}{c_{\mathrm{W}}^2}\right)}{2u^2} \\
&+ \frac{3c_{\mathrm{W}}^2\left(16c_{\mathrm{W}}^2+19u\right)\left(\ln\left(c_{\mathrm{W}}^2\right)-\ln(u)\right)}{2u} \\
&+ \frac{3\left(4c_{\mathrm{W}}^4 - 50c_{\mathrm{W}}^2 u + 19u^2\right)\Phi\left(\sqrt{u},c_{\mathrm{W}},c_{\mathrm{W}}\right)}{2\left(4c_{\mathrm{W}}^2-u\right)u}, \qquad \text{(I.35)}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{F}_3(u,\omega) =& \frac{\alpha_{em}^2}{576\,\pi^2\,c_{\mathrm{W}}^4 s_{\mathrm{W}}^4}\,\frac{m_\mu^2}{M_Z^2}\,\Bigg[ \frac{9u\left(2c_{\mathrm{W}}^2 - u + \omega\right)}{\omega} \\
&+ \left[A_1(u,\omega)\ln\left(\frac{u}{c_{\mathrm{W}}^2}\right) + 9c_{\mathrm{W}}^4\left(c_{\mathrm{W}}^4 - 4c_{\mathrm{W}}^2\omega + 3\omega^2\right)\ln(c_{\mathrm{W}}^2)\right]\frac{\ln\left(\omega/c_{\mathrm{W}}^2\right)}{2\omega^2(c_{\mathrm{W}}^2-\omega)} \\
&+ A_2(u,\omega)\frac{\ln\left(u\right)}{\omega(4c_{\mathrm{W}}^2-u)} + A_3(u,\omega)\frac{\ln\left(\omega\right)}{\omega(c_{\mathrm{W}}^2-\omega)} \\
&+ A_4(u,\omega)\frac{\ln\left(c_{\mathrm{W}}^2\right)}{\omega^2(4c_{\mathrm{W}}^2-u)(c_{\mathrm{W}}^2-\omega)} + \frac{A_5(u,\omega)}{c_{\mathrm{W}}^2\omega^2}\mathrm{Li}_2\left(1-\frac{u}{c_{\mathrm{W}}^2}\right) \\
&+ \frac{A_6(u,\omega)}{uc_{\mathrm{W}}^2(4c_{\mathrm{W}}^2-u)^2(c_{\mathrm{W}}^2-\omega)}\Phi\left(\sqrt{u},c_{\mathrm{W}},c_{\mathrm{W}}\right) \\
&+ \frac{A_7(u,\omega)}{\omega^2\left(c_{\mathrm{W}}^2-\omega\right)\left(c_{\mathrm{W}}^4 - 2c_{\mathrm{W}}^2(u+\omega) + (u-\omega)^2\right)}\Phi\left(\sqrt{u},\sqrt{\omega},c_{\mathrm{W}}\right)\Bigg], \qquad \text{(I.36)}
\end{aligned}
$$

$$
\begin{aligned}
A_1(u,\omega) =& -9c_{\mathrm{W}}^2u^3+9c_{\mathrm{W}}^2u^2\left(3c_{\mathrm{W}}^2+\omega\right)+27c_{\mathrm{W}}^4u\left(\omega-c_{\mathrm{W}}^2\right)\\
&+9\left(c_{\mathrm{W}}^8-4c_{\mathrm{W}}^6\omega+3c_{\mathrm{W}}^4\omega^2\right), \qquad \text{(I.37)}\\
A_2(u,\omega) =& \frac{9c_{\mathrm{W}}^4\omega}{2}-9u^2\left(5c_{\mathrm{W}}^2+\omega\right)+u\left(36c_{\mathrm{W}}^4+\frac{153c_{\mathrm{W}}^2\omega}{4}\right)+9u^3, \qquad \text{(I.38)}\\
A_3(u,\omega) =& 9c_{\mathrm{W}}^2u^2-\frac{9}{2}c_{\mathrm{W}}^2u\left(4c_{\mathrm{W}}^2+\omega\right), \qquad \text{(I.39)}\\
A_4(u,\omega) =& -\frac{9}{2}u^2\omega\left(2c_{\mathrm{W}}^4+9c_{\mathrm{W}}^2\omega+2\omega^2\right)\\
&+\frac{9}{8}u\omega\left(32c_{\mathrm{W}}^6+13c_{\mathrm{W}}^4\omega+35c_{\mathrm{W}}^2\omega^2\right)+9u^3\omega^2, \qquad \text{(I.40)}\\
A_5(u,\omega) =& -9u^3\left(c_{\mathrm{W}}^2+\omega\right)-9u\left(3c_{\mathrm{W}}^6+2c_{\mathrm{W}}^2\omega^2\right)\\
&+9u^2\left(3c_{\mathrm{W}}^4+4c_{\mathrm{W}}^2\omega+\omega^2\right)+\frac{9}{2}c_{\mathrm{W}}^4\left(2c_{\mathrm{W}}^4-6c_{\mathrm{W}}^2\omega+\omega^2\right), \qquad \text{(I.41)}\\
A_6(u,\omega) =& -9u^4\left(9c_{\mathrm{W}}^2+\omega\right)+u\left(81c_{\mathrm{W}}^6\omega-225c_{\mathrm{W}}^8\right)+9c_{\mathrm{W}}^8\left(\omega-c_{\mathrm{W}}^2\right)\\
&-\frac{9}{2}u^2\left(3c_{\mathrm{W}}^6+37c_{\mathrm{W}}^4\omega\right)+u^3\left(198c_{\mathrm{W}}^4+72c_{\mathrm{W}}^2\omega\right)+9u^5, \qquad \text{(I.42)}\\
A_7(u,\omega) =& -9c_{\mathrm{W}}^2u^4+18c_{\mathrm{W}}^2u^3\left(2c_{\mathrm{W}}^2+\omega\right)+36u\left(c_{\mathrm{W}}^8-2c_{\mathrm{W}}^6\omega\right)\\
&-9c_{\mathrm{W}}^2u^2\left(6c_{\mathrm{W}}^4-c_{\mathrm{W}}^2\omega+\omega^2\right)-9c_{\mathrm{W}}^2\left(c_{\mathrm{W}}^2-3\omega\right)\left(c_{\mathrm{W}}^3-c_{\mathrm{W}}\omega\right)^2. \qquad \text{(I.43)}
\end{aligned}
$$

# Bibliography


[1] B. Aubert et al. [BABAR Collaboration], Phys. Rev. Lett. 92, 121801(2004) "Mesonium and anti-mesonium". Zh. Eksp. Teor. Fiz. 33 (2): 549–551. February 1957. reproduced and translated in B. Pontecorvo (February 1957). "Mesonium and Antimesonium". Sov. Phys. JETP. 6 (2): 429–431. Bibcode:1958JETP....6..429P.

[2] B. Pontecorvo (May 1968). "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". Zh. Eksp. Teor. Fiz. 53: 1717–1725.

[3] Davis, Raymond; Harmer, Don S.; Hoffman, Kenneth C. (1968). "Search for Neutrinos from the Sun". Physical Review Letters. 20 (21): 1205–1209. Bibcode:1968PhRvL..20.1205D. doi:10.1103/PhysRevLett.20.1205.

[4] Ahmad, Q.R.; Allen, R.C.; Andersen, T.C.; Anglin, J.D.; Buhler, G.; Barton, J.C.; et al. (SNO Collaboration) (2001-07-25). "Measurement of the rate of $\nu_e + d \rightarrow p + pe^-$ interactions produced by 8B Solar neutrinos at the Sudbury Neutrino Observatory". Physical Review Letters. 87 (7): 071301. arXiv:nucl-ex/0106015.

[5] Fukuda, Y.; et al. (Super-Kamiokande Collaboration) (24 August 1998). "Evidence for Oscillation of Atmospheric Neutrinos". Physical Review Letters. 81 (8): 1562–1567. arXiv:hep-ex/9807003

[6] A. D. Sakharov (1967). "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe". Journal of Experimental and Theoretical Physics Letters. 5: 24–27.

[7] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, "Evidence for the $2\pi$ Decay of the $K_2^0$ Meson," Phys. Rev. Lett. 13, 138 (1964)

[8] Trimble, V. (1987). "Existence and nature of dark matter in the universe". Annual Review of Astronomy and Astrophysics. 25: 425–472.

[9] Allen, Steven W.; Evrard, August E.; Mantz, Adam B. (2011). "Cosmological Parameters from Clusters of Galaxies". Annual Review of Astronomy and Astrophysics. 49 (1): 409–470. arXiv:1103.4829

[10] Peebles, P. J. E.; Ratra, Bharat (2003). "The cosmological constant and dark energy". Reviews of Modern Physics. 75 (2): 559–606. arXiv:astro-ph/0207347

[11] J. Ellis, "Limits of the Standard Model", arXiv:hep-ph/0211168

[12] Kolmogorov, A. N. , (2017), "Foundations of the Theory of Probability", 1933.

[13] Cox, R. T. (1946). "Probability, Frequency and Reasonable Expectation". American Journal of Physics. 14: 1–10

[14] Kass, R.E., Raftery, A.E., "Bayes Factors", *Journal of the American Statistical Association*, 1995, Vol.90, No.430. pp.773-795

[15] Cousins, Robert D. "Why isn't every physicist a Bayesian?" *American Journal of Physics* **63** (5) pp 398-410, May 1995

[16] Rickmers, A.D., Todd, H.N. *Statistics An Introduction*

[17] K. A. Olive et al. [Particle Data Group Collaboration], "Review of Particle Physics," Chin. Phys. C 38 (2014) 090001

[18] M. Davier, A. Hoecker, B. Malaescu and Z. Zhang, Eur. Phys. J. C 71 (2011) 1515 [Erratum-ibid. C 72 (2012) 1874] [arXiv:1010.4180 [hep- ph]]; arXiv:1706.09436 [hep-ph].

[19] A. Keshavarzi, T. Teubner, talks at "g-2 theory initiative," June 2017, Fermilab, and at "PhiPsi 2017", June 2017, Mainz.

[20] F. Jegerlehner, arXiv:1705.00263 [hep-ph].

[21] A. Chapelain, "The Muon g-2 experiment at Fermilab", arXiv:1701.02807

[22] The Super-Kamiokande Collaboration, Y. Fukuda et al, "Evidence for oscillation of atmospheric neutrinos", arXiv:9807003

[23] Kersten, J., Park, J., Stockinger, D., Velasco-Sevilla, L., "Understanding the correlation between $(g-2)_\mu$ and $\mu \to e\gamma$ in the MSSM", arXiv:1405.2972

[24] MEG Collaboration: J. Adam, X. Bai, A. M. Baldini et al, "New constraint on the existence of the $\mu^+ \to e^+\gamma$" decay", arXiv:1303.0754

[25] Bailer-Jones, Coryn A. L., (2017), "Practical Bayesian Inference", Cambridge University Press.

[26] Brod, J., Lenz, A., Tetlalmatzi-Xolocotzi, G., Wiebusch, M."New physics effects in tree-level decays and the precision in the determination of the CKM angle $\gamma$", arXiv:1412.1446

[27] W. Altmannshofer, D. M. Straub,"New physics in $b \to s$ transitions after LHC run 1", arXiv:1411.3161 [hep-ph].

[28] A. Lenz, "Lifetimes and HQE", arXiv:1405.3601 [hep-ph].

[29] Y. Amhis et al., (Heavy Flavor Averaging Group (HFAG), "Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012", arXiv:1207.1158 [hep-ex].

[30] A. Lenz, U. Nierste, "Numerical updates of lifetimes and mixing parameters of B mesons", arXiv:1102.4274 [hep-ph].

[31] V. M. Abazov et al. [D0 Collaboration], Phys. Rev. D 86 (2012) 072009 [arXiv:1208.5813 [hep-ex]].

[32] J. P. Lees et al. [BaBar Collaboration], Phys. Rev. Lett. 111 (2013) 10, 101802 [Addendum-ibid. 111 (2013) 15, 159901] [arXiv:1305.1575 [hep-ex]].

[33] R. Aaij et al. [LHCb Collaboration], arXiv:1409.8586 [hep-ex].

[34] J. P. Lees et al. [ The BABAR Collaboration], arXiv:1411.1842 [hep-ex].

[35] V. M. Abazov et al. [D0 Collaboration], Phys. Rev. Lett. 110 (2013) 011801 [arXiv:1207.1769 [hep-ex]].

[36] R. Aaij et al. [LHCb Collaboration], Phys. Lett. B 728 (2014) 607 [arXiv:1308.1048 [hep-ex]].

[37] J. D. Bjorken. Nucl. Phys. Proc. Suppl. 11, 325 (1989)

[38] P.Ball, "$|V_{ub}|$ from UTangles and $B \to \pi l\nu$", arXiv:hep-ph/0611108.

[39] B. Aubert, et al, (BABAR Collaboration), "Measurement of the Time-Dependent CP Asymmetry in $B_0 \to D(*)^0_{CP} h_0$ Decays".

[40] Bobeth, C. et al, "On New Physics in $\Delta\Gamma_d$", arXiv:1404.2531

[41] M. Ibe, A. Rajaraman, Z. Surujon, "Does Supersymmetry Require Two Higgs Doublets?", arXiv:1012.5099 [hep-ph]

[42] Drees, M., Godbole, R., Roy, P., (2005), "Theory and Phenomenology of Sparticles", World Scientific Publishing.

[43] T. Moroi, "Muon anomalous magnetic dipole moment in the minimal supersymmetric standard model", Phys. Rev. D53 (1996) 6565–6575, arXiv:hep-ph/9512396. Erratum ibid. D56 (1997) 4424.

[44] W. Altmannshofer, D.M. Straub, "Viability of MSSM scenarios at very large $\tan\beta$"arXiv:1004.1993 [hep-ph]

[45] D. Stockinger, "The Muon Magnetic Moment and Supersymmetry", J. Phys. G 34 (2007) R45–R92, arXiv:hep-ph/0609168.

[46] Raby, S. "Yukawa unification : MSSM at large $\tan\beta$", arXiv:1210.2693 [hep-ph].

[47] A. Brignole, A. Rossi, "Anatomy and Phenomenology of $\mu - \tau$ Lepton Flavour Violation in the MSSM", arXiv:hep-ph/0404211/

[48] Y. Omura, et al. "$\tau-$ and $\mu-$physics in a general two Higgs doublet model with $\mu - \tau$ flavour violation", Phys Rev D (94), 055019 (2016).

[49] K. Hayasaka, K. Inami, Y. Miyazaki, K. Arinstein, V. Aulchenko, et al., "Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced $\tau^+\tau^-$ Pairs," Phys.Lett., vol. B687, pp. 139–143, 2010, 1001.3221, arXiv:0711.2189.

[50] Panico, Giuliano, Riembau, Marc, Vantalon, Thibaud. (2017). "Probing light top partners with CP violation." Journal of High Energy Physics. 2018. 10.1007/JHEP06(2018)056.

[51] Cherchiglia, A., Kneschke, P., Stockinger, D., Stockinger-Kim, H., "The muon magnetic moment in the 2HDM: complete two-loop result", arXiv:1607.06292.

[52] D. Aristizabal Sierra, A. Vicente, "Explaining the CMS Higgs flavor violating decay excess", arXiv:1409.7690 [hep-ph].

[53] S. M. Barr and A. Zee, Phys. Rev. Lett. 65 (1990) 21 [Erratum-ibid. 65 (1990) 2920].

[54] V. Ilisie,"New Barr-Zee contributions to $(g-2)_\mu$ in two-Higgs-doublet models".

[55] B. e. Lautrup, A. Peterman and E. de Rafael, "The anomalous magnetic moment of the muon and short-distance behaviour of quantum electrodynamics",Phys. Rept. 3 (1972) 193.

[56] J. P. Leveille,"he second-order weak correction to $(g-2)$ of the muon in arbitrary gauge models", Nucl. Phys. B 137 (1978) 63.

[57] A. Dedes and H. E. Haber, "Can the Higgs sector contribute significantly to the muon anomalous magnetic moment?" JHEP 0105 (2001) 006 [hep-ph/0102297].

[58] A. Cherchiglia, D. Stockinger, H. Stöckinger-Kim, Phys. Rev. D 98, 035001 (2018) arXiv:1711.11567 [hep-ph]

[59] D. Lopez-Val and J. Sola, Eur. Phys. J. C 73, 2393 (2013) [arXiv:1211.0311 [hep-ph]]

[60] D. Chang, W. F. Chang, C. H. Chou and W. Y. Keung, Phys. Rev. D 63 (2001) 091301 doi:10.1103/PhysRevD.63.091301 [hep-ph/0009292].

[61] K. m. Cheung, C. H. Chou and O. C. W. Kong, Phys. Rev. D 64 (2001) 111301 doi:10.1103/PhysRevD.64.111301 [hep-ph/0103183].

[62] D. Chang, W. S. Hou and W. Y. Keung, "Two loop contributions of flavor changing neutral Higgs bosons to $\mu \to e\gamma$," Phys. Rev. D 48, 217 (1993) [hep-ph/9302267].

[63] Talk by Pierre Savard in European Physical Society Conference on High Energy Physics 2015, 27 July 2015, Vienna, Austria; The ATLAS Collaboration,"Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector", arXiv:1508.03372 [hep-ex].

[64] V. Khachatryan et al. [CMS Collaboration],"Search for lepton-flavour-violating de- cays of the Higgs boson," arXiv:1502.07400 [hep-ex].

[65] L. Calibbi, G. Signorelli, "Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction", [arXiv:1709.00294]

[66] M. Blanke et al., JHEP 5, 013 (2007).

[67] C.-X. Yue and Sh. Zhao, Eur. Phys. J. C 50, 897 (2007).

[68] Si. Karmakar, S. Rakshit, "Relaxed constraints on the heavy scalar masses in 2HDM",Phys. Rev. D 100, 055016.

[69] Y. Omura, E. Senaha and K. Tobe, "Lepton-flavor-violating Higgs decay $h \to \mu\tau$ and muon anomalous magnetic moment in a general two Higgs doublet model," JHEP 1505, 028 (2015) [arXiv:1502.07824 [hep-ph]].

[70] H. E. Logan, D. MacLennan, "Charged Higgs phenomenology in the flipped two Higgs doublet model", arXiv:1002.4916 [hep-ph]

[71] A. Celis, V. Ilisie and A. Pich, JHEP 1312, 095 (2013) [arXiv:1310.7941 [hep-ph]].

[72] J. Cho, T. Y. Kim, J. Song, "Constraints from Heavy Higgs boson masses in the two Higgs doublet model", arXiv:1801.09514 [hep-ph]

[73] K.G. Hayes et al. (Mark II Collaboration), "Experimental Upper Limits on Branching Fractions for Unexpected Decay Modes of the Tau Lepton", Phys. Rev. D 25, 2869 (1982).

[74] Athron, P., et al, "FlexibleSUSY 2.0: Extensions to investigate the phenomenology of SUSY and non-SUSY models", arXiv:1710.03760

[75] Li, Tianjun, "Determining $M_1$ , $M_2$ , $\mu$ and $tan(\beta)$ in MSSM from Chargino and Neutralino Masses", hep-ph/9810242

[76] Omura, Y., Senaha, E., Tobe, K., "$\tau-$ and $\mu-$ physics in a general two Higgs doublet model with $\mu-\tau$ flavour violation", Phys Rev D 94, 044019 (2016)

[77] A. Crivellin, A. Kokulu and C. Greub,"Flavor-phenomenology of two-Higgs-doublet models with generic Yukawa structure," Phys. Rev. D 87, no. 9, 094031 (2013) [arXiv:1303.5877 [hep-ph]].

[78] A. Broggio, E. J. Chun, M. Passera, K. M. Patel and S. K. Vempati, JHEP 1411 (2014) 058 [arXiv:1409.3199 [hep-ph]].

[79] E. J. Chun and J. Kim, JHEP 1607, 110 (2016) [arXiv:1605.06298 [hep-ph]].

[80] L. Wang and X. F. Han, JHEP 1505, 039 (2015) [arXiv:1412.4874 [hep-ph]].

[81] T. Abe, R. Sato and K. Yagyu, JHEP 1507, 064 (2015) [arXiv:1504.07059 [hep-ph]].

[82] A. Crivellin, J. Heeck and P. Stoffer, Phys. Rev. Lett. 116 (2016) no.8, 081801 [arXiv:1507.07567 [hep-ph]].

[83] E. J. Chun, Z. Kang, M. Takeuchi and Y. L. S. Tsai, JHEP 1511 (2015) 099 [arXiv:1507.08067 [hep-ph]].

JHEP **1504** (2015) 077 [arXiv:1502.04199 [hep-ph]].

[84] A. I. Davydychev and J. B. Tausk, "Two loop selfenergy diagrams with different masses and the momentum expansion", Nucl. Phys. B **397**, 123 (1993).

[85] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher and J. P. Silva, Phys. Rept. 516, 1 (2012) [arXiv:1106.0034 [hep-ph]]

[86] Z. Chacko and G. D. Kribs, "Constraints on lepton flavor violation in the minimal supersymmetric standard model from the muon anomalous magnetic moment measurement", Phys. Rev. D64 (2001) 075015, arXiv:hep-ph/0104317.

[87] F. Krinner, A. Lenz, T. Rauh, "The inclusive decay b to ccs revisited", arXiv:1305.5390 [hep-ph]

[88] M. Beneke, G. Buchalla, C. Greub, A. Lenz, U. Nierste, "Next-to-Leading Order QCD Corrections to the Lifetime Difference of $B_s$ Mesons", arXiv:hep-ph/9808385.

[89] M. Beneke, G. Buchalla, A. Lenz, U. Nierste, CP asymmetry in flavour-specific B decays beyond leading logarithms", arXiv:hep-ph/0307344.

[90] A. Lenz, U. Nierste, "Theoretical update of $B_s - \bar{B}_s$ mixing", arXiv:hep-ph/0612167/

[91] G. Bell, V. Pilipp, "$B \to \pi - \pi_0/\rho - \rho_0$ to NNLO in QCD factorization", arXiv:0907.1016 [hep-ph]

[92] A. Lenz, "Selected Topics in Heavy Flavour Physics", arXiv:1404.6197 [hep-ph].

[93] Robert Piessens, Elise deDoncker-Kapenga, Christian Ueberhuber, David Kahaner, QUADPACK: A Subroutine Package for Automatic Integration, Springer, 1983, ISBN: 3540125531, LC: QA299.3.Q36.

[94] T. Han, S. K. Kang and J. Sayre, "Muon $g-2$ in the Aligned Two Higgs Doublet Model" ,JHEP 1602, 097 (2016), [arXiv:1511.05162 [hep-ph]].

[95] https://arxiv.org/pdf/1911.05543.pdf